

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:56:37; Search time 2599 Seconds

(without alignments)  
5307.707 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474  
Sequence: 1 gccacgattatcagtcgacgagtaaa 474

Scoring table: IDENTITY NUC  
Gapop 10-0, Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

1 number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: GenBank

- 1: gb\_ba:\*
- 2: gb\_hcg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_hcg\_hum:\*
- 31: em\_hcg\_in:\*
- 32: em\_hcg\_om:\*
- 33: em\_hcg\_ov:\*
- 34: em\_hcg\_pat:\*
- 35: em\_hcg\_ph:\*
- 36: em\_hcg\_pl:\*
- 37: em\_hcg\_pr:\*
- 38: em\_hcg\_ro:\*
- 39: em\_hcg\_sts:\*
- 40: em\_hcg\_un:\*
- 41: em\_hcg\_vi:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474	100.0	474	6	AX088165 Sequence
2	439.2	92.7	1191	6	AX128507 Sequence
3	439.2	92.7	2274	9	AF190500 Sequence
4	222	46.8	187431	9	AF190500 Homo sapi
5	171.8	36.2	1068	6	AX148176 Sequence
6	171.8	36.2	2049	6	AX385037 Sequence
7	171.8	36.2	2142	6	AX451566 Sequence
8	171.8	36.2	2193	6	AX385032 Sequence
9	171.8	36.2	2214	6	AX451562 Sequence
10	171.8	36.2	2265	6	AX385027 Sequence
11	171.8	36.2	2436	9	AF453828 Homo sapi
12	171.8	36.2	2838	9	AF453828 Homo sapi
13	170.6	36.0	58377	2	AC128964 Sequence
14	170.6	36.0	124282	2	AC121415 Sequence
15	170.6	36.0	173546	2	AC098607 Rattns no
16	163	34.4	179055	2	AC122462 Mus muscu
17	156.8	33.1	321	6	AX384516 Sequence
18	147.2	31.1	173546	2	AC098607 Rattns no
19	146.2	30.8	2214	6	AX385045 Sequence
20	146.2	30.8	2539	10	AF346501 Mus muscu
21	106.8	22.5	1018	6	AX147820 Sequence
22	106.8	22.5	170522	9	AL138708 Human DNA
23	92.4	19.5	180673	10	AC068627 Mus Muscu
24	92.4	19.5	213462	10	AC077689 Mus Muscu
25	86	18.1	164610	2	AC098990 Rattns no
26	86	18.1	171869	2	AC121029 Rattns no
27	61.4	13.0	176565	2	AL136106 Homo sapi
28	51.6	10.9	171869	2	AC121029 Rattns no
29	48.8	10.3	1015	6	AX254453 Sequence
30	47	9.9	104992	2	AC005504 Plasmodu
31	47	9.9	169546	2	AC004157 Plasmodu
32	45.6	9.6	122258	2	AL356427 Homo sapi
33	45.6	9.6	133751	9	HS1131C10 Human DNA
34	44.4	9.4	174186	9	AC026394 Homo sapi
35	44	9.3	5516	9	AF090694 Homo sapi
36	44	9.3	5644	9	AF090694 Homo sapi
37	44	9.3	6905	9	AF090693 Homo sapi
38	44	9.3	78359	9	AF314199S67 Homo sapi
39	44	9.3	177829	9	AC026392 Homo sapi
40	44	9.3	181507	2	AC015907 Homo sapi
41	44	9.3	187495	9	AC026887 Homo sapi
42	43	9.1	30493	9	AL662819 Human DNA
43	42.4	8.9	153903	9	AC008380 Homo sapi
44	42.4	8.9	153906	9	AC008503 Homo sapi
45	42.4	8.9	170921	9	AC025030 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	SEQUENCE	1 from Patent WO0114548.	DNA	linear	PAT 17-MAR-2001
AX088165	AX088165	Sequence 1	AX088165	AX088165	AX088165.1	GI:13397078	human.					

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
New g-protein coupled receptor and dna sequences thereof  
Patent: WO 0114548-A 1 01-MAR-2001;

FEATURES MERCK PATENT GmbH (DE)  
Location/Qualifiers  
1..474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

CDS  
1..474  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC34686.1"  
/db\_xref="GI:13397079"  
/translation="AQLSVAFPLGINLAFLIIVFSYGMFVHQSALITAEIRNO  
VKEMILARFRPIFTDLCMTPIFVAKPSLLOVEIPGTTSMVIVGSAINSAIN  
PLIYLTITPPEKMHFRFWMNTYRKSMDSKIRKMLHSSGCGCHCRCHLS"

BASE COUNT 144 a 92 c 88 g 150 t  
ORIGIN

Query Match 100.0%; Score 474; DB 6; Length 474;  
Best Local Similarity 100.0%; Pred. No. 4.3e-100; Indels 0; Gaps 0;  
Matches 474; Conservative 0; Mismatches 0;

1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTGGCCGATTTATCATATA 60  
Db 1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTGGCCGATTTATCATATA 60  
QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGCCATAACAGAACTGAA 120  
Db 61 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGCCATAACAGAACTGAA 120  
QY 121 ATACGGAATCAAGTTAATAAAGAGATGATCCTTGCCAAAGTTTCTTATAGTATT 180  
Db 121 ATACGGAATCAAGTTAATAAAGAGATGATCCTTGCCAAAGTTTCTTATAGTATT 180  
QY 181 ACTGATGATTAATGCTGGATACCCATTTTGTAGCGAAACCTCTTCACCTTCAGTGA 240  
Db 181 ACTGATGATTAATGCTGGATACCCATTTTGTAGCGAAACCTCTTCACCTTCAGTGA 240  
QY 241 GAATACCAAGTACCATTAACCTCTTGAGTGTATGTTATTCGATTAACAGTGT 300  
Db 241 GAATACCAAGTACCATTAACCTCTTGAGTGTATGTTATTCGATTAACAGTGT 300  
QY 301 TTGAACCAATTTCTTAACTCTGACCAACAAGCATTTTAAAGAAATGATTCATCGGTTT 360  
Db 301 TTGAACCAATTTCTTAACTCTGACCAACAAGCATTTTAAAGAAATGATTCATCGGTTT 360  
QY 361 TGGCATTAATCTAGACCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 420  
Db 361 TGGCATTAATCTAGACCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 420  
QY 421 CATCATTCATCTGGGGGGAATGTGGCCACTGACAGAGATGCCACTGAGTTAA 474  
Db 421 CATCATTCATCTGGGGGGAATGTGGCCACTGACAGAGATGCCACTGAGTTAA 474

RESULT 2  
AX128507 1191 bp DNA linear PAT 15-MAY-2001  
LOCUS AX128507  
DEFINITION Sequence 15 from Patent WO0131014.  
ACCESSION AX128507  
VERSION AX128507.1 GI:14134974  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1191)  
Vogeli, G., Wood, L.S. and Merchant, K.  
G protein-coupled receptors expressed in brain  
JOURNAL Patent: WO 0131014-A, 15 03-MAY-2001;  
PHARMACIA & UPJOHN COMPANY (US)  
FEATURES  
1..1191  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

CDS  
1..1191  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC38938.1"  
/db\_xref="GI:14134975"  
/translation="MFRPLVNLSHLYEKKFOYGVAPHVRSCKPNTDGISSLLENLLAS  
IIOIRFVWVAVTCFNGFIIVICMRPIYSEKLVYMSIISLCCADLMGIVLFLVGG  
FDLKPGEYKNAOLMMESTHOOLVSLILSTREVSLLLTLETKYICIVYPRCV  
RQKCRITTVLLITWTGTIVAPIPLSNKEFPKNYGNVCGVCEPLHSEDTESIGAOIY  
SVAIFPLGINLAFLIIVFSYGMFVHQSALITAEIRNOVKEMILARFRPIFTD  
ALCMIPFVVKFPLSLQVEIPGTTSMVIVFLPINSALNPLIYLTITPPEKMHFR  
WNTYRKSMDSKQKTYVAFSIWEMWPELQEMPELMDLFTYPCESMLISQSTRL  
NSYS"

BASE COUNT 340 a 229 c 226 g 396 t  
ORIGIN

Query Match 92.7%; Score 439.2; DB 6; Length 1191;  
Best Local Similarity 96.8%; Pred. No. 4.7e-92;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTGGCCGATTTATCATATA 60  
Db 643 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTGGCCGATTTATCATATA 702  
QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGCCATAACAGAACTGAA 120  
Db 703 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGTGCCATAACAGAACTGAA 120  
QY 121 ATACGGAATCAAGTTAATAAAGAGATGATCCTTGCCAAAGTTTCTTATAGTATT 180  
Db 763 ATACGGAATCAAGTTAATAAAGAGATGATCCTTGCCAAAGTTTCTTATAGTATT 180  
QY 181 ACTGATGATTAATGCTGGATACCCATTTTGTAGCGAAACCTCTTCACCTTCAGTGA 240  
Db 823 ACTGATGATTAATGCTGGATACCCATTTTGTAGCGAAACCTCTTCACCTTCAGTGA 240  
QY 241 GAATACCAAGTACCATTAACCTCTTGAGTGTATGTTATTCGATTAACAGTGT 300  
Db 883 GAATACCAAGTACCATTAACCTCTTGAGTGTATGTTATTCGATTAACAGTGT 300  
QY 301 TTGAACCAATTTCTTAACTCTGACCAACAAGCATTTTAAAGAAATGATTCATCGGTTT 360  
Db 943 TTGAACCAATTTCTTAACTCTGACCAACAAGCATTTTAAAGAAATGATTCATCGGTTT 360  
QY 361 TGGCATTAATCTAGACCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 420  
Db 1003 TGGCATTAATCTAGACCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 420  
QY 421 CATCATTCATCTGGGGGGAATGTGGCCACTGACAGAGATGCCACTGAGTTAA 474  
Db 1061 CATCATTCATCTGGGGGGAATGTGGCCACTGACAGAGATGCCACTGAGTTAA 474

\* RESULT 3  
AF190500 2274 bp mRNA linear PRI 10-AUG-2001  
LOCUS AF190500  
DEFINITION Homo sapiens leucine-rich repeat-containing G protein-coupled  
receptor 7 (LGR7) mRNA, complete cds.  
ACCESSION AF190500  
VERSION AF190500.1 GI:10441729  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 2274)  
Hsu, S.Y., Kudo, M., Chen, T., Nakabayashi, K., Bhalla, A., van der  
Spek, P.J., van Duijn, M. and Hsueh, A.U.  
The three subfamilies of leucine-rich repeat-containing G  
protein-coupled receptors (LGR): identification of LGR6 and LGR7  
and the signaling mechanism for LGR7  
JOURNAL Mol. Endocrinol. 14 (8), 1257-1271 (2000)  
MEDLINE 20388592

PUBMED 10935549 2 (bases 1 to 2274)  
REFERENCE Hsu, S.Y. and Hsueh, A.J.W.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB 5385, Stanford, CA 94305, USA

FEATURES  
source  
1..2274  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..2274  
/gene="LGR7"  
1..2274  
/gene="LGR7"  
/codon\_start=1  
/product="leucine-rich repeat-containing G protein-coupled receptor-7"  
/protein\_id="AA017167.1"  
/db\_xref="gi:10441730"  
/translation="MTSGSVFFYLLIPKYPFHGGGDVKSGLGYPFGCINTKCLPOL  
LHNGVDVDCQAOEDNCGDNGNMGMODKFPASYIKMTSGYPEAPETBETLGSVPR  
CLOGLLEDCDEFTNIRAPVSVSNATMRSGVNLIRLDPDCRYTHDLKLYLQNN  
KITSITYAPRGASLITKLYSHNRITFLKGFVEDLHRLFWLIIEDNHSLISPPFE  
YGLNSLILVLMNVNVLRLPDKPLCOHPRHLWLDLGNHITHNLRTFLSCSNLTPL  
VMRNKLNHNLNENTFAPLOKLELDLDSNKNLEMLPPLIFDLKELSLQNLSSYNLQKI  
QANPDYLVKLSLSLEGEISNLIQGMFRLNMLSHIFKKFQYCGGAPVRSCREN  
TDGLSLENLFIASIGLEISNLIQGMFRLNMLSHIFKKFQYCGGAPVRSCREN  
CLMGILYFPRCVARPKGCRITVYLLIMINGFIYAFPLSKKEFKNYGTNGCFPL  
EKYICIVPRCVARPKGCRITVYLLIMINGFIYAFPLSKKEFKNYGTNGCFPL  
HSYEPSEISGQIYVSALPLGLNLAFLIYVSIGSMYSVHQSATATETIRNOVKEM  
ILARFPIYFTDLQWPIFVVPKLSLOVEIGTTSWVPIPLINSALNDILYTT  
LTPNFKEMHFWNYRNRKSDMSKQKTYAPSPFVEMWVPLEMEPELKKPDLFTY  
PCENSLISQSHRLMSYS"

BASE COUNT 669 a 460 c 421 g 724 t  
ORIGIN

Query Match 92.7%; Score 439.2; DB 9; Length 2274;  
Best Local Similarity 96.8%; Pred. No. 4,3e-92;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTTCTGGATTAATTGGCCGCAATTATCATCATTA 60  
Db 1726 GCCCAGATTATTCAGTGGCAATTTTCTGGATTAATTGGCCGCAATTATCATCATTA 1785

QY 61 GTTTTTCCTATGAGAGCATGTTTATATGTTTCATCAAGAGTCCATAACAGCAACTGAA 120  
Db 1786 GTTTTTCCTATGAGAGCATGTTTATATGTTTCATCAAGAGTCCATAACAGCAACTGAA 1845

QY 121 ATACGAGATCAAGTTAAAGAGAGATGATCCCTGGCAACGTTTTCTTATATGATATT 180  
Db 1846 ATACGAGATCAAGTTAAAGAGAGATGATCCCTGGCAACGTTTTCTTATATGATATT 1905

QY 181 ACTGATGATTATGCTGGATAACCAATTTTGTAGGAAACCTCTTCACTGCTTCAGGTA 240  
Db 1906 ACTGATGATTATGCTGGATACCAATTTTGTAGGAAATTTCTTTCACCTGCTTCAGGTA 1965

QY 241 GAAATPACAGGTATACATTAACCTCTTGGGTAGTGGTATGTTATTTCTGCCATTAACAGTCT 300  
Db 1966 GAAATPACAGGTATACATTAACCTCTTGGGTAGTGGTATGTTATTTCTGCCATTAACAGTCT 2025

QY 301 TTGAACCAATTTCTATACTCTGACCAACAGACCATTTAAGAAATGATTCATCGGTTT 360  
Db 2026 TTGAACCAATTTCTATACTCTGACCAACAGACCATTTAAGAAATGATTCATCGGTTT 2085

QY 361 TTGGGATACTACACAAAGAAATCTATGACACAGCAAGGATATCAGAAACATATGCTC 420  
Db 2086 TTGGGATACTACACAAAGAAATCTATGACACAGCAAGG-TCAGAAACATATGCTC 2143

QY 421 CATCATTCATCTGGGGGAAATGTGGCCATCGCAGAGAGATGCCACTGAGTTAA 474  
Db 2144 CATCATTCATCTGGGGGAAATGTGGCCATCGCAGAGAGATGCCACTGAGTTAA 2197

```

RESULT 4
AC107219          187431 bp      DNA      linear      PRI 27-MAR-2002
LOCUS
DEFINITION Homo sapiens BAC clone RP11-575B4 from 4, complete sequence.
ACCESSION AC107219
VERSION AC107219.5 GI:19526153
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens.
            Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PubMed 9847074
REFERENCE
AUTHORS 2 (bases 1 to 187431)
TITLE Isak, A., Meyer, R. and Boyer, E.
JOURNAL The sequence of Homo sapiens BAC clone RP11-575B4
REFERENCE
AUTHORS 3 (bases 1 to 187431)
TITLE Unpublished (2001)
JOURNAL
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE
AUTHORS 4 (bases 1 to 187431)
TITLE Waterston, R.H.
JOURNAL Direct Submission
            Submitted (21-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE
AUTHORS 5 (bases 1 to 187431)
TITLE Waterston, R.H.
JOURNAL Direct Submission
            Submitted (16-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE
AUTHORS 6 (bases 1 to 187431)
TITLE Waterston, R.
JOURNAL Direct Submission
            Submitted (27-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Mar 16, 2002 this sequence version replaced gi:1885160.
COMMENT
            -----
            Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@watson.wustl.edu
            -----
            Summary Statistics
            Center project name: H_NH0575B04
            -----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this

```

SOURCE INFORMATION:  
The Rpct-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Fritgen, E., Tatem, M., Catanesse, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is CTD-2005A22; the clone sequenced to the right is Rp11-663M18. Actual start of this clone is at base position 1 of Rp11-575B4; actual end is at base position 187431 of Rp11-575B4.

Sequence derived from one plasmid subclone, base position 102011 to 102054.

```
FEATURES
source
    Location/Qualifiers
        1..187431 Homo sapiens
        /db_xref="taxon:9606"
        /chromosome="4"
        /map="4"
        /clone="Rp11-575B4"
        /clone_1b="Rpct-11"
        1..408
        /rpt_family="L1"
        repeat_region
            563..697
            /rpt_family="Alu"
        repeat_region
            829..938
            /rpt_family="Alu"
        repeat_region
            1227..1326
            /rpt_family="MIR"
        repeat_region
            1327..2602
            /rpt_family="L1"
        repeat_region
            2973..3009
            /rpt_family="AT_rich"
        repeat_region
            3233..3534
            /rpt_family="Alu"
        repeat_region
            3582..3616
            /rpt_family="AT_rich"
        repeat_region
            3685..3772
            /rpt_family="L2"
        repeat_region
            4141..4297
            /rpt_family="MAlR"
        repeat_region
            4298..4593
            /rpt_family="Alu"
        repeat_region
            4594..4804
            /rpt_family="MAlR"
        repeat_region
            5020..5436
            /rpt_family="MAlR"
        repeat_region
            5483..5506
            /rpt_family="AT_rich"
        repeat_region
            6006..6170
            /rpt_family="MAlR"
        repeat_region
            6709..6748
            /rpt_family="L2"
        repeat_region
            6863..7446
            /rpt_family="L1"
        repeat_region
            7454..8078
            /rpt_family="ERV1"
        repeat_region
            8079..8175
            /rpt_family="L1"
        repeat_region
            8176..8477
            /rpt_family="Alu"
        repeat_region
            8176..8189
            /note="match to EST A1683385 (NID:g4893567) tx66609.x1"
        repeat_region
            8478..8745
            /rpt_family="L1"
        repeat_region
            8746..9052
            /rpt_family="Alu"
        repeat_region
            9053..9124
```

```
repeat_region
    /rpt_family="L1"
    9126..9148
    /rpt_family="(TTTT)n"
misc_feature
    9681..9926
    /note="similar to Mus musculus EST BB297224 (NID:g16401965)"
repeat_region
    9921..10022
    /rpt_family="L1"
misc_feature
    10225..10665
    /note="similar to Mus musculus EST BB297224 (NID:g16401965)"
misc_feature
    10870..11388
    /note="match to EST A1653269 (NID:g4737248) wb44e10.x1"
misc_feature
    11115..11388
    /note="match to EST A1652121 (NID:g4736100) wb47c02.x1"
repeat_region
    12295..12442
    /rpt_family="Alu"
repeat_region
    12443..12734
    /rpt_family="Alu"
repeat_region
    12735..12900
    /rpt_family="Alu"
repeat_region
    13540..13819
    /rpt_family="Alu"
repeat_region
    14177..14344
    /rpt_family="Alu"
repeat_region
    14358..15112
    /rpt_family="L1"
repeat_region
    15130..15158
    /rpt_family="(TA)n"
repeat_region
    15194..15487
    /rpt_family="Alu"
repeat_region
    15575..15853
    /rpt_family="Alu"
misc_feature
    16606..16891
    /note="match to EST A1375172 (NID:g4175162) tc10b12.x1"
misc_feature
    16789..17233
    /note="match to EST A1805323 (NID:g5391889) tc32h05.x1"
repeat_region
    17004..17094
    /rpt_family="L2"
misc_feature
    complement(17219..17220)
    /note="match to EST AA122078 (NID:g1678116) zk93g10.r1"
misc_feature
    17228..17513
    /note="match to EST AA122078 (NID:g1678116) zk93g10.r1"
repeat_region
    17508..17621
    /rpt_family="L1"
misc_feature
    17610..18078
    /note="match to EST A1089319 (NID:g3428378) qb16d05.x1"
repeat_region
    17655..17710
    /rpt_family="L1"
misc_feature
    17849..18074
    /note="match to EST AA626188 (NID:g2538575) zv89a01.s1"
repeat_region
    18230..18295
    /rpt_family="Alu"
repeat_region
    18298..18349
    /rpt_family="Alu"
repeat_region
    18425..18699
```

```
Query Match
Best Local Similarity 97.8%; Score 222; DB 9; Length 187431;
Matches 225; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 29 TTGGTATTAAATTTGGCCGATTTATCATGATTTTCTTATGGAAGCATGTTTATA 88
DB 13203 TAGGATTTAATTTGGCCGATTTATCATGATTTTCTTATGGAAGCATGTTTATA 13262
QY 89 GGTTCATCAAAAGTGCATACAGCACTGAATTCGGAATCAAGTTAAAAAGAGATGA 148
DB 13263 GGTTCATCAAAAGTGCATACAGCACTGAATTCGGAATCAAGTTAAAAAGAGATGA 13322
QY 149 TCGTTCGCAAAAGTTCCTTATGATTTATGATGATCAATTAAGCTGATACCCATT 208
DB 13323 TCGTTCGCAAAAGTTCCTTATGATTTATGATGATCAATTAAGCTGATACCCATT 13382
```



QY	209	TTGTACCGAAACCTCTTTCACGTCTTCAGTGTGAATACAGGTACATA	258
Db	13383	TTGTAGTGAATTTCTTTCACGTCTTCAGGTGAATACAGGTACATA	13432
RESULT 5			
LOCUS	AX148176	1068 bp	DNA
DEFINITION	Sequence 17 from Patent WO0136471.		linear
ACCESSION	AX148176		
VERSION	AX148176.1		
KEYWORDS	GI:14347081		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 1068)		
FEATURES	Chen, R., Dang, H. T., and Lowitz, K. P.		
FEATURES	Endogenous and non-endogenous versions of human g protein-coupled		
FEATURES	receptors		
FEATURES	Patent: WO 0136471-A 17 25-MAY-2001;		
FEATURES	Arena Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
FEATURES	1..1068		
FEATURES	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
BASE COUNT	278 a 215 c 214 g 361 t		
ORIGIN			
Query Match	36.2%; Score 171.8; DB 6; Length 1068;		
Best Local Similarity	67.4%; Pred. No. 7e-30;		
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;			
QY	10	TATTCAGTGGCAATTTTCTTGGATTAATTTATTTGGCCGCAATTATCATGATGTTTTTC	69
Db	568	TATTCCTTGGAAATTTCTCTAGTGGAACCTTGCGGCTTTCTCATGTTGTGTTTCC	627
QY	70	TATGGAACGATGTTTATTAAGTGTATCAAAAGGCCATPAACGAAATACGAAAT	129
Db	628	TATATTTACTAGTGTGTTGTTCCATTTAAAAAACCGGCTTGACAGCCACGAAATGAAGAT	687
QY	130	CAAGTTAAAAAGAGATATCCCTTSCCAAAGCTTTTCTTAAATGATTTACAGATCA	189
Db	688	TGTTTGAAGAGAGGTGGCTGTGCAATATGTTTCTTTTATAGTGTCTCTGATCC	747
QY	190	TTATGCTGGAATGCCATTTTGTAGCGAAACCTCTTTCAGTCTCAGTGAATATCA	249
Db	748	ATTCGCTGGATTCCTGATTTGTAGTTAAATCCCTTCCCTCCGGGGAATATCA	807
QY	250	GGTACCATPAACCTCTGGGTGTGATTTGGTATATGTCATTAACAGCTTTGAACCA	309
Db	808	GACACAAATGACTCTCGTATATGATATTTTTCCTTCAGATTAACAGCTTTGAATCA	867
QY	310	ATTCTTAACTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTTGGCATTA	368
Db	868	ATCCTTAACTCTCACAAACCACTTTTAAAGCAAGTTGAAGAACAGCTCTGACAA	926
RESULT 6			
LOCUS	AX385037	2049 bp	DNA
DEFINITION	Sequence 11 from Patent WO0214489.		linear
ACCESSION	AX385037		
VERSION	AX385037.1		
KEYWORDS	GI:19578160		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1		
TITLE	Paszty, C. J., Gong, J., Daugherty, B., and Rogers, N.		
TITLE	Lauche-rich repeat-containing g-protein coupled receptor-8		
TITLE	molecules and uses thereof		

JOURNAL		Patent: WO 0214489-A 11 21-FEB-2002;	
FEATURES		Amgen, Inc. (US)	
Source	Location/Qualifiers		
1..2049	/organism="Homo sapiens"		
/db_xref="taxon:9606"			
1..2049	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD28590.1"		
	/db_xref="GI:19578161"		
	/translation="MIVFLVFKHLFSLRLTMFELHPIVLINVKDFALTOGSMITPPE COKGFFPCGNLTKLCIPRAFHCDGDGDCNAGDEBNCQGSMAITFGVHGANSVAL TOEGLFOYPOCCDCKEETELFECNAGDLSKSVMI SNMVTLSLKNKHISLPDKYIKRX TKLKKIPLQHNCHIRHISKAPKPGYCNOLITLLDNPTRISORLPFGINSLEPILSMV NYLAPKQWCAQWQPOLNMDLBNRKRITKYLNSPLFSDSLTVLDDSSNITTELSPH FKDKLQKDLBERLEIPNITRMFOEPKMSHLYFKNFRCSYAPHVICMPLTIDG SSEFBLNANLRLIPWVIAITCTGMLFVIGMSFIAENVTHAMSILKILCADCLAL GVYLFVGIIPDKIRGOYKXALMBSVQCRINGFLMSTBSVLILTYLLEKFP VIVPEFSNIRGKQTSVILICIMWAGPLIIVIPMNVDYGFNFGKNGVCEPLYDD TEDPFSKYSGLIFLWNLAFILVIFSIWAGPLIIVIPMCSIOKTALOTTVEVRNFGREAVARA REFPVSFSDAICMIIPVFPVKTLSPVSIPTDMSWIVIPLPVNSALNPILYTLTMM PFKXKQGLAHKQRKSIPIKIKKSLSLSIWIEDSSLKIGVINKITLIGDSIMKPVSA "		
BASE COUNT	579 a 415 c 395 g 660 t		
ORIGIN			
Query Match	36.2% Score 171.8; DB 6; Length 2049;		
Best Local Similarity	67.4%; Pred. No. 6.4e-30;		
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;			
QY 10	TATTCAGTGGCAATTTTCTTGGTATTAATTTGGCGGATTTATCATATGTTTTC 69		
Db 1549	TATTCCTTGGAAATTTCTAGTGTGAACCTTGCGCTTTCTCATCATGTGTTTTCC 1608		
QY 70	TATGAAGCATGTTTATAGTGTTCATCAAGTGGCATAAACGCACTGAATACGGAAT 129		
Db 1609	TATATTACTATGTTCTGTTCATTCATAAAACCGGCTTGCAAGACCAAGAAATAGGAAT 1668		
QY 130	CAAGTAAAAAAGATGATATCTCTGCCAAAGTGTTCCTTATATGTAATTAAGTATGA 189		
Db 1669	TGTTTGAAGAGAGGTGGCGTTGCAATGCTTTCTTTTATAGTGTCTCTATGCC 1728		
QY 190	TTATGCTGATTAACCATTTTGTGAGCGAACCCTCTTCACTGCTTCAGTGAATAACCA 249		
Db 1729	ATCGCTGATTCCTCGTATTTGTGATTAATAATCCCTTCCTCCGCGGTGAATACCA 1788		
QY 250	GGTACCATTAACCTCTGGGTGATGTTGGTATTCGTGCATTAACAGTCTTGAACCA 309		
Db 1789	GACACATGACTCTCTGATATGTAATTTTTTCTTCCAGTTAAACAGTCTTGAATCA 1848		
QY 310	ATTCTCTACTCTGACCAAGACCATTTTAAAGAAATGATTCATCGGTTTTGGCATTA 368		
Db 1849	ATCCCTACTACTCTCACAAACCACTTTTAAAGGACAAAGTTGAACACAGCTGTCACAA 1907		
RESULT 7			
LOCUS	AX451566 2142 bp DNA linear PAT 03-JUL-2002		
DEFINITION	Sequence 5 from Patent WO0226824.		
ACCESSION	AX451566		
VERSION	AX451566.1 GI:21698550		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Booker, J. N., Mintier, G., Ramnathan, C. S. and Hawken, D. R.		
AUTHORS	homo; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	Feder, J. N., Mintier, G., Ramnathan, C. S. and Hawken, D. R.		
	A novel human g-protein coupled receptor, hgsr1bmy5, expressed		
	highly in brain and ovarian tissues		
	Patent: WO 0226824-A 5 04-APR-2002;		

FEATURES	Bristol Myers Squibb Company (US)
source	location/qualifiers
	1..2142
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	613 a 435 c 408 g 686 t
ORIGIN	

Query Match	36.2%	Score 171.8	DB 6	Length 2142
Best Local Similarity	67.4%	Pred. No. 6.3e-30		
Matches 242; Conservative	0	Mismatches 117		Indels 0; Gaps 0;

Oy	10	TATTCAGGCGCAATTTTCTTGGAATTAATTGGCCGCAATTATCACTCATTTATTTTCC	69
Db	1642	TATTCCTTGGAATTTTCTTAGGGTGMACTTCGTGGCTTTTCTCATCAATGTGTATTC	1701
Oy	70	TATGAGACAGTTTATTAAGTGTTCATCAAGAGGCCATMAAGCAACATGAAATPAGGAAT	129
Db	1702	TATATTAATAAGTGTCTGTCCATTCAAAAAACCCTTGACACACAGAAATAGGAAT	1761
Db	130	CAAGTTAAAAAAGAGATGATCCCTGCCAAACGTTTTTCTTTATAGTATTTACTGATCA	189
Db	1762	TGTTTGGAAAGAGGGGCGCTGTGCAAAACGTTTTCTTTTAAAGTGTCTGTAGTGC	1821
Oy	190	TTATGCTGGATACCCATTTTGTGAGCAACCTCTTCACTGCTTAGGTAATAATATCA	249
Db	1822	ATTCGCTGGATTCCTGTATTTGTATTAATCCCTTCCCTCTTCCGGGTGAANAATCA	1881
Oy	250	GGTACCATTAACCTCTTGGGTATGATTTGTTATCTGCACCTTAACAGTCTTTGAACCA	309
Db	1882	GACAAATGACTTCTCGATATGATATTTTTTCCCTTCAGTTAACAGTGCTTGAATCA	1941
Oy	310	ATTCTCTACTCTGACCAACAAGACCATTTAAGAANAATGATATGCGGTTTGGCATTA	368
Db	1942	ATCCCTCTACTCTCACAAACCACTTTTTTAAAGACAAAGTTGAACAAGCTGCTGCACAA	2000

RESULT 8  
 AX385032 2193 bp DNA PAT 19-MAR-2002  
 LOCUS  
 DEFINITION Sequence 6 from Patent WO0214489.  
 ACCESSION AX385032  
 VERSION AX385032.1 GI:19578156  
 KEYWORDS  
 SOURCE .  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Paszty, C.J., Gong, J., Daugherty, B. and Rogers, N.  
 Leucine-rich repeat-containing g-protein coupled receptor-8  
 molecules and uses thereof  
 Patent: WO 0214489-A 6 21-FEB-2002;  
 Amgen, Inc. (US)  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 CDS  
 1..2193  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2193  
 /note="unamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD28588.1"  
 /db\_xref="GI:19578157"  
 /translation="MIVPLVFVKHLFSLITLIMFFLHFTIVLINVKDFALDQGSWITPS  
 CQKGFPCGNLTKCLPRAFHCDGKDCGAGDEBNCSDTSMATIFGVHGNANSLAL  
 TQECFLQYQCCDCKETELCEVNDLKSVPISNNVLTLSKOKHSLEDPKFIYK  
 TKIKTIFLOHNCIRHISRAKPAFGCLNLIYLVNHNITTLPGILFKDLHQLTWILIDD  
 NITRIQSRLFTGINSIFLSPWNNVLEALPKWCAQWQDNPVLDENGRNIXYLTNSTP  
 FLSCSITVLDSNMTITELSPHLFKDLKLOKNTLSNPMLYHKNQFESLKOQLSL  
 DLERFIEIPNINRMQCPMKNTSHIYFKNRFQCSYAPHYRICMPLTIDGISSEDLANNK  
 ILIRFWVIAFTTCFQNLFTVIGMRSFITAEHTTHAMSIKILCCADCLMGVYLFVGVLE  
 DIKRFQYQYKVALMMWESVQCRMLQFLAMLSLEVSALTLLTYLTLEKFLVIVFPFSNIR  
 PKRQTSVLICIMWAGFLIAYIPFMNDYFQNFYGNKGQVFLIYQTDIEDISGYS

sig peptide 1..108  
 LGIFLGUNLAFILIVFSYITWMCSDLOKTLQTEVNRNGREVAANRPPIVFSDA  
 ICMIPFVVKILSLFRVEIDPTMTSVIFLFEVSNALNPIYTLTTFPKDKLQLL  
 HKHQRKSIFFIKRKSLTSLTWIEDSSSLGLVINKLITLQDSIMKPS"  
 BASE COUNT 626 a 456 c 411 g 700 t  
 ORIGIN

Query Match	36.2%	Score 171.8	DB 6	Length 2193
Best Local Similarity	67.4%	Pred. No. 6.3e-30		
Matches 242; Conservative	0	Mismatches 117	Indels 0	Gaps 0

Qy	10	TATTCAGTGGCAATTTTCTTGATATAATTTGGCGCAATTAATATATCAATAGTTTTTCC	69
Db	1693	TATTCCTTGGAATTTTCCTAGGTGGAATCTTGCGCTTTCTCATATGTGTGTTTCC	1752
Qy	70	TATGGAAGCATGTTTTATATCTGTTCATCAAGTGCATTAACGCAACGTAAATACGAAAT	129
Db	1753	TATATTTCTATGTTCGTGTTCCATTTCAAAAAACGCTTTCACAGCCAGAGATAGAAAT	1812
Qy	130	CAAGTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTATATGATTTTACTGATGCA	189
Db	1813	TGTTTTGGAGAGAGGTGGCTGTGGCAATCGTTTTCTTTTATATGTTCTCTGATGCG	1872
Qy	190	TTATGCTGGAATACCATTTTTTTGTATGCGAAACCTCTTTCACGTCTCAGGTAAATATACCA	249
Db	1873	ATTCGCTGGATTCCTGTATTTGTATGATTAATCTTTTCCCTCTTCGGGTGAAATATACCA	1933
Qy	250	GGATACATTAACCTTTGGGTAGATGTGTTATTTCTGCATTAAACAGTCTTTGAAACCCA	309
Db	1933	GACACAAATGATTCCTCTGGATATGATTTTTTTTCCCTTCAGTAAACAGTCTTTGGAATCCA	1993
Qy	310	ATTCTCTATATCTCTGACCACACAGACATTTAAAGAAATATTCATACGCTTTGGCATTA	368
Db	1993	ATCCCTATATCTTCCACACCAACTTTTTTAAAGACAAGTTGAACAGCTGCTGCACAA	2051

	AX451562		2214 bp	DNA	linear	PAT 03-JUL-2002
LOCUS	AX451562					
DEFINITION	Sequence 1 from Patent WO0226824.					
ACCESSION	AX451562					
VERSION	AX451562.1	GI:21698547				
KEYWORDS	.					
SOURCE	human.					
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1	Feder J.N., Mintier G., Ramanathan C.S. and Hawken D.R. A novel human g-protein coupled receptor, hghrmy5, expressed highly in brain and ovarian tissues Patent: WO 0226824-A 1 04-APR-2002;				
AUTHORS	Bristol-Myers Squibb Company (US)					
TITLE	Location/Qualifiers					
JOURNAL	1..2214					
FEATURES	/organism="Homo sapiens" /db_xref="taxon:9606"					
source	BASE COUNT 638 a 457 c 416 g 703 t					
ORIGIN	Query Match 36.2%; Score 171.8; DB 6; Length 2214; Best Local Similarity 67.4%; Pred No.6.3e-30; Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;					
QY	10 TATTGAGTGCAATTTTCTGTGGTAATAATTGGCGCATTTATCATCATAAGTTTTTC 69					
Db	1714 TATTCTCTTGGAAITTTTCTTAGGTGAACCTGGCGCTTTCTCATCATGTGTGTTTCC 1773					
QY	70 TATGAGAAGATGTTTATAGTGTTCATCAAAGTGCCATACAGCAACTGAATTAGGAAT 129					
Db	1774 TATATTACTATGTGTTGTTCCAATTCMAAAAACGCCCTTCAGACCACAGAAAGTAAGAAAT 1833					
QY	130 CAAGTAAAAAAGATGATCCTTGCACAACGTTTTTCTTTATATAGTATTTACTGATGA 189					

[illegible]

Db	1885	TGTTTGGAGAGAGGTGCCTGTGGAAATCGTTTTCTTTTATATAGTGTCCTGATGCC	1944
Oy	190	TTATGCTGATACCCATTGTTTGGAGGAACCCTTCACTGCTTAGGATAGAATAACCA	249
Db	1945	ATCTGCTGATTCCTGATATTGTTAGTTAATAATCCTTCCCTCTCCGGGATGAATACCA	2004
Oy	250	GGTACCATTAACCTCTTGGGTAGTAGTGGTTATTTCTGCATTAAACAGTCTTGAACCCA	309
Db	2005	GACACAATGACTTCCTGGATAGTAGATTTTTCCTCCCTTCAGTTAAGAGTGCTTGAATCCA	2064
Oy	310	ATTCTTATACCTTGACCAAGAACATTTAAAGAATATGATTCATGCTTTGGCATAA	368
Db	2065	ATCCTCTATACCTCAACAACTTTTAAAGACAGATTGAACAGCTGCTGCACAA	2123
RESULT 11			
AF453828			
LOCUS	AF453828	2436 bp	mRNA linear PRI 04-FEB-2001
DEFINITION	Homo sapiens G protein-coupled receptor affecting testicular		
ACCESSION	AF453828		
VERSION	AF453828.1	GI:18483167	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 2436)		
JOURNAL	Gorlov,I.P., Kamet,A., Jones,E., Lamb,D., Truong,A., Bogatcheva,N.,		
REFERENCE	Bishop,C.E., McElreavey,K. and Agoulnik,A.I.		
AUTHORS	Mutations of the GREAT gene cause cryptorchidism		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 2436)		
FEATURES	Agoulnik,A.I.		
source	Direct Submission		
	Submitted (30-NOV-2001) Ob/Gyn, Baylor College of Medicine, 6550		
	Fannin St., Su. 861, Houston, TX 77030, USA		
	Location/Qualifiers		
	1..2436		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="13"		
	/map="13q12-q13"		
	1..2436		
	/gene="GREAT"		
	71..2335		
	/gene="GREAT"		
	/codon_start=1		
	/product="G protein-coupled receptor affecting testicular		
	descent"		
	/protein_id="AAI73946.1"		
	/db_xref="gi:18483168"		
	/translation="MIVFLVKHLFSLRLTMFLPHLFIILNVKDFALIQGSMTTPS		
	CQKYRFGGNITKCLPEAFHCDDGDGNGDENBDCGTSMAITFGVHNANSVALP		
	TOEFLKQPQCCKDETELECVNADGDKSVPMISNNVTLLSKKKHSLSDEYIKRY		
	TKLKIFLHNCIHISRKAFFGLCNALILNMCNTTLTRPGI FKDIHLTWLITLDI		
	NPIRISORLPFTGNLSIPLSMANNYALEIPKQCAOMPQNWDLEBNRIKYLTNS		
	PLSDSITVLFLENQIGFPEPKTFSSKNIAGEIDLSNITTELSPHFQDLLOQT		
	MLSNPLMYLHKQNFESLKQLQSLDRLEIEIPINTWPFQPMKNLSHYFNPRYS		
	APHRIEMPLETHDGISSFEDLLANNILRIFWVAVFICFGNLPTYGMSFKANTTT		
	AMSFKILCADCLMGVLFEVGIDPIKRGQYQXALLMBSVOCRLGFAMISTEV		
	VULLTVLLEKFLVIPPEPSNIRPGKRSVILICIMMGFLAVIPFMWKDYEGNRN		
	YGRKGCFRELVAYDQTEDIGSKCYSLGILEGVNLAPLIIVSYITMCSIOKTLQOTT		
	ENVRGCFRELVAYNRRFFLYPSDAICMIPIPVFKILSIFRVEIDPTMSVIIFLPP		
	NSAIPLFLYLITLNFFFDKLQKLLHKQRKSIFIKIKSLSTSIWILEDSSKLGVV		
	NKTLTSGINSIMPVS"		
BASE COUNT	703 a	503 c	456 g 774 t
ORIGIN			
Query Match	36.2%	Score 171.8;	DB 9; Length 2436;
Best Local Similarity	67.4%;	Pred. No. 6.2e-30;	
Matches 242; Conservative	0;	Mismatches 117;	Indels 0; Gaps 0;

QY 10 TATTCAGTGCATATTTTCTGTGTAATTAATTTGGCCCATTTATCATCATAGTTTTC 69  
DB 1835 TATTCCTGTGGAATTTTCTGTGTAATTAATTTGGCCCATTTATCATCATAGTTTTC 1894  
QY 70 TATGGAAGCATGTTTATGTCATCAAGTCCCAATTAAGCACTGAATACGAAT 129  
DB 1895 TATATTAATGATGTCATCAAGTCCCAATTAAGCACTGAATACGAAT 1954  
QY 130 CAAGTAAAGAGATGATCCTTGCAAAAGTTTCTTATAGTAATTAATGATGCA 189  
DB 1955 TGTTTGGAAAGAGGTCGTTGCAAAAGTTTCTTATAGTAATGATGATGCC 2014  
QY 190 TTAATGTCGATACCATTTTGTAGCGAAACCTTTTCACTGCTTCAAGTAATCA 249  
DB 2015 ATCTGTCGATCTCTGTAATTTGTAGTAATTAATCTTCTTCCGCGGTGAATCA 2074  
QY 250 GGTACCAATACCTCTGGGTAGTGTATGTTATCTGCAATTAAGTGTGCAATCA 309  
DB 2075 GACACAAATGACTTCTGATAGTATGTTTCTTCCATGATTAAGTGTGCAATCA 2134  
QY 310 ATCTCTATCTGACCAAGACCATTTTAAGAAATGATTCGCTTTGGCATTA 368  
DB 2135 ATCTCTATCTGACCAAGACCATTTTAAAGCAAGTTGAAGCAAGTGTGCAATCA 2193  
RESULT 12  
AF403384 2838 bp mRNA linear PRI 19-FEB-2002  
LOCUS Homo sapiens LGR8 mRNA, complete cds.  
DEFINITION AF403384  
ACCESSION AF403384.2 GI:18702459  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 2838)  
Heu, S.-Y., Nakabayashi, K., Nishi, S., Kumagai, J., Kudo, M.,  
Sherwood, O.D. and Heu, A.J.  
Activation of orphan receptors by the hormone relaxin  
Science 295 (5555), 671-674 (2002)  
TITLE  
JOURNAL  
PUBMED  
11809971  
REFERENCE  
AUTHORS  
2 (bases 1 to 2838)  
Heu, S.-Y., Nakabayashi, K. and Bhalla, A.  
Direct Submision  
Submitted (26-JUL-2001) GYN/OB, Stanford University, MSOB S385,  
Stanford, CA 94305, USA  
3 (bases 1 to 2838)  
Heu, A.J.W.  
Direct Submision  
Submitted (19-FEB-2002) Division of Reproductive Biology,  
Department of Gynecology and Obstetrics, Stanford University  
Medical Center, 300 Pasteur Drive, Room A-344, Stanford, CA  
94305-5317, USA  
REMARK  
COMMENT  
On Feb 19, 2002 this sequence version replaced gi:18419431.  
FEATURES  
source  
1..2838  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
/map="2371"  
/codon\_start=1  
/product="LGR8"  
/protein\_id="AA169324.2"  
/db\_xref="GI:18702459"  
/db\_xref="GI:18702460"  
/translation="MIVFVFKDHLPSLRITMFLPIYLIVKDALVQSGMTTS  
CQKGFPCGNTLKLPAFHLVDKDDCGNADENCGDSGMATFETVGNANVAL  
TOBCLKQYPOCCCKETELCVAGDLSKVMISNNVTLLSLKKNKTHSLPKVFIKY  
TKLKKIFQHNCRHISKAFKFLCNLIIVLNHCITTLRPGIFKDLHQLTWLILD

NTPIRISORLFTGLNSLPLFSLMVNNVLEALPKOMCAQPOLANWVDEGNRIKYLNST  
FLGSDSLTVLPBPNOIGFPEKTFESLKLGLDLSNNTITELSPHFLKLOL  
NLSSNPMTYHKNRQFESLKOLOSIDLERITIPNTPMOPMNSHIVKPNRCSY  
APHVRICMPLTIDGISFEDLLANNILITFWVAIFITFCNLUVIGRSIRKXNTH  
AMSIRKICADCLMGVLPFVGVGIDIKRQYKRYALMWSVQCRIMGLAMSTVB  
SVLLITLLEKREKPLVIVFPSPNIRPGKROSVLILCMMAGFLIAIIPFNKQYFGNF  
YKRGVCEPLYDQTEDIDIGSKYSLGIFLVNLLAFILIVFSLITMFCSLOKTALQTT  
EVRNCPGRVAVANRPFPIVFPSPNIRPGKROSVLILCMMAGFLIAIIPFNKQYFGNF  
NSALNPILYLTITNFPDKLKLQHLHQRKRSIRKIKKSLSTSIWIEDSSILGLVL  
NKITLIGDSIMKEVNS"  
BASE COUNT 817 a 616 c 536 g 869 t  
ORIGIN  
Query Match 36.2% Score 171.8; DB 9; Length 2838;  
Best Local Similarity 67.4%; Pred. No. 6.1e-30;  
Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
QY 10 TATTCAGTGCATATTTTCTGTGTAATTAATTTGGCCCATTTATCATCATAGTTTTC 69  
DB 1871 TATTCCTGTGGAATTTTCTGTGTAATTAATTTGGCCCATTTATCATCATAGTTTTC 1930  
QY 70 TATGGAAGCATGTTTATGTCATCAAGTCCCAATTAAGCACTGAATACGAAT 129  
DB 1931 TATATTAATGATGTCATCAAGTCCCAATTAAGCACTGAATACGAAT 1990  
QY 130 CAAGTAAAGAGATGATCCTTGCAAAAGTTTCTTATAGTAATTAATGATGCA 189  
DB 1991 TGTTTGGAAAGAGGTCGTTGCAAAAGTTTCTTATAGTAATGATGATGCC 2050  
QY 190 TTAATGTCGATACCATTTTGTAGCGAAACCTTTTCACTGCTTCAAGTAATCA 249  
DB 2051 ATCTGTCGATCTCTGTAATTTGTAGTAATTAATCTTCTTCCGCGGTGAATCA 2110  
QY 250 GGTACCAATACCTCTGGGTAGTGTATGTTATCTGCAATTAAGTGTGCAATCA 309  
DB 2111 GACACAAATGACTTCTGATAGTATGTTTCTTCCATGATTAAGTGTGCAATCA 2170  
QY 310 ATCTCTATCTGACCAAGACCATTTTAAGAAATGATTCGCTTTGGCATTA 368  
DB 2171 ATCTCTATCTGACCAAGACCATTTTAAAGCAAGTTGAAGCAAGTGTGCAATCA 2229  
RESULT 13  
AC128964 58377 bp DNA linear HTG 24-JUN-2002  
LOCUS Rattus norvegicus clone CH230-17915, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION  
AC128964  
AC128964.1 GI:21953739  
ACCESSION  
VERSION  
HTG: HTGS PHASE1.  
KEYWORDS  
Rattus norvegicus.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus  
1 (bases 1 to 58377)  
REFERENCE  
AUTHORS  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Albrooks, S.L., Amaralunge, H.C., Are, J.R., Ayala, M., Banks, T.,  
Barbata, J., Benton, U., Blinige, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowles, S., Brieva, M., Brown, M., Brown, M., Bryant, N.F.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, M., Cavan, C.,  
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dim, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Geo, J., Garcia, A., Garner, T., Garza, N., Gail, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Homsj, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C.,  
Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, J., Liu, J., Liu, W., Louised, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenko, S., Ogun, M., Okunnu, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, U., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojudo, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shoochari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansy, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 58377)  
Worley, K.C.  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KBQJ  
Center clone name: CH230-17915  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 20312 bases at least Q40  
Consensus quality: 22274 bases at least Q30  
Consensus quality: 23493 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 36 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1034: contig of 1034 bp in length  
\* 1035 1134: gap of unknown length  
\* 1135 2575: contig of 1441 bp in length  
\* 2576 2675: gap of unknown length  
\* 2676 4009: contig of 1334 bp in length  
\* 4010 4109: gap of unknown length  
\* 4110 5591: contig of 1482 bp in length  
\* 5592 5691: gap of unknown length  
\* 5692 6877: contig of 1186 bp in length  
\* 6878 6977: gap of unknown length  
\* 6978 8557: contig of 1580 bp in length  
\* 8558 8657: gap of unknown length  
\* 8658 9929: contig of 1271 bp in length  
\* 9929 10028: gap of unknown length  
\* 10029 11169: contig of 1141 bp in length  
\* 11170 11269: gap of unknown length

11270 12351: contig of 1082 bp in length  
\* 12352 12451: gap of unknown length  
\* 12452 12493: contig of 1042 bp in length  
\* 13494 13593: gap of unknown length  
\* 13594 13988: contig of 1395 bp in length  
\* 14989 15088: gap of unknown length  
\* 15089 16389: contig of 1301 bp in length  
\* 16390 16489: gap of unknown length  
\* 16490 17697: contig of 1208 bp in length  
\* 17698 17797: gap of unknown length  
\* 17798 18057: contig of 1260 bp in length  
\* 19058 19157: gap of unknown length  
\* 19158 20422: contig of 1265 bp in length  
\* 20423 20522: gap of unknown length  
\* 20523 21549: contig of 1027 bp in length  
\* 21550 21649: gap of unknown length  
\* 21650 22953: contig of 1304 bp in length  
\* 22954 23053: gap of unknown length  
\* 23054 24690: contig of 1637 bp in length  
\* 24691 24790: gap of unknown length  
\* 24791 25985: contig of 1195 bp in length  
\* 25986 26085: gap of unknown length  
\* 26086 27287: contig of 1202 bp in length  
\* 27288 27387: gap of unknown length  
\* 27388 29766: contig of 2379 bp in length  
\* 29767 29866: gap of unknown length  
\* 29867 31067: contig of 1201 bp in length  
\* 31068 31167: gap of unknown length  
\* 31168 32710: contig of 1543 bp in length  
\* 32711 34287: gap of unknown length  
\* 34288 34387: contig of 1477 bp in length  
\* 34388 36393: gap of unknown length  
\* 36394 36493: contig of 2006 bp in length  
\* 36494 37830: gap of unknown length  
\* 37831 37930: contig of 1337 bp in length  
\* 37931 39971: gap of unknown length  
\* 39972 40071: contig of 2041 bp in length  
\* 40072 41200: gap of unknown length  
\* 41201 41300: contig of 1129 bp in length  
\* 41301 42420: gap of unknown length  
\* 42421 44250: contig of 1120 bp in length  
\* 44251 44723: gap of unknown length  
\* 44724 44823: contig of 2203 bp in length  
\* 44824 46668: gap of unknown length  
\* 46669 46768: contig of 1845 bp in length  
\* 46769 49052: gap of unknown length  
\* 49053 49152: contig of 2284 bp in length  
\* 49153 50929: gap of unknown length  
\* 50930 51029: contig of 1777 bp in length  
\* 51030 52705: gap of unknown length  
\* 52706 52806: contig of 1677 bp in length  
\* 52807 55072: gap of unknown length  
\* 55073 55172: contig of 2266 bp in length  
\* 55173 58377: gap of unknown length  
\* 58378 58377: contig of 3205 bp in length.

FEATURES  
source 1. 58377  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-17915"

BASE COUNT 14489 a 12592 c 12867 g 12553 t 5876 others  
ORIGIN

Query Match 36.0%; Score 170.6; DB 2; Length 58377;  
Best Local Similarity 81.7%; Pred. No. 7.5e-30;  
Matches 197; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Oy 31 GGTATTATTTGGCGGCGATTATCATCATAGTTTTCCTATGGAACGATTTTATAGT 90  
|||||  
DB 47183 GGTATTACCTGGGGCGGCTTATCATCATTTGTCTTCATGGAACGATTTTACAGT 47242  
|||||  
Oy 91 GTTCATCAAGTGCATTAACGCAACTGAATAACGATCAAGTTAAAAAGATGATC 150  
|||||

OY		47243	GTCATCAAGACACATAAAGCACAACCGAAATACAGAAGCACGTGAAGAAGAGATGATGC	47302
OY	151	CTTGCCAAACGGTTTTTTTCTTTTAAGATTACTGATGACATTATGCTGGATACCATTITTT	210	
Db	47303	CTCGCCAAAACGCTTTTCTTTTATTTGTTCCACCGCATGCACTTGTGTGGATTCCTCATTTTTT	47362	
OY	211	GTAGGAAACCGCTTTCACCTGCTTAGGTAAGAAAAATACACGAGTACCATAAACCTCTTGCGGTA	270	
Db	47363	AATAGTAAATTTCTCTCACTAGTATGGGGTGGAAATPACAGTACAACTTTTCTTGAGAA	47422	
OY	271 G 271			
Db	47423	G 47423		
RESULT 14 ACI21415/c LOCUS DEFINITION MISSION VERSION KEYWORDS SOURCE ORGANISM				
REFERENCE AUTHORS				
Murphy D.M., Adams C., Adio-Oduola B., Ali-Isman F.R., Allen C., Aldebrooks S.L., Amaraturung H.C., Are J.R., Ayelle M., Banks T., Barbato A.J., Benton J., Blinage K., Blankenburg K., Bonini D., Bouck J., Bowle S., Briveva M., Brown E., Brown M., Bryant N.P., Buluy C., Burich P., Burkett C., Butrell K.L., Byrd N.C., Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C., Coyle M.D., Datborne S.R., David R., Davila M.L., Davis C., Day-Carroll L., Dedrich D.A., Delaney K.R., Delgado O., Denn A.U., Ding Y., Dinhl H.H., Doughwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J., Earhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M., Falls T., Ferraguto D., Flagg N., Ford U., Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R., Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K., Harris K., Harris K., Hart M., Havlak P., Hawes A., Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B., Homel F., Howard S., Huber J., Huylk S., Hume J., Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolyvet S., Jouhad S., Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovac C., Kratovic U., Kureshi A., Landry N., Leal B., Lewis L.C., Lewig L., Li J., Li Z., Lichtarge O., Lieu C., Liu J., Lu W., Louiseau H., Lozodo R.J., Lu X., Lucier A., Lucier R., Luna A., Ma J., Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E., Massey E., Mashiney E., McLeod M.P., Meador M., Mei G., Melcher M., Miner G., Minner Z., Mitchell T., Monabbak K., Morgan M., Morris S., Moser M., Neal D., Newton J., Newtson N., Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nmekewko S., Ogih M., Okunou G., Orangunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L., Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y., Rivers M., Rojks A., Rojnubokhan I., Rolfe M., Ruiz S., Savary G., Scherer S., Scott G., Shen H., Shoocharit N., Sibson I., Sodergren E., Sonaike T., Sparks A., Stanley H., Stone K., Sutcon A., Svatek A., Tabour P., Tamerasia A., Tamerasia K., Tang H., Taney J., Taylor C., Taylor T., Telofod B., Thomas N., Thomas S., Usmani K., Vaquerre L., Vera V., Villalon D., Vinson R., Wang Q., Wang S., Ward-Moree S., Warren R., Washington C., Watlington S., Williams G., Williamson A., Wislaczky R., Wooden S., Worley K., Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D., Weinstein G. and Gibbs R.				
TITLE JOURNAL REFERENCE AUTHORS				
Unpublished 2 (bases 1 to 124282)				
Worley,K.C.				

REFERENCE JOURNAL	REFERENCE JOURNAL	COMMENT
<p>TITLE JOURNAL</p> <p>Submitted (18-May-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>3 (bases 1 to 124282)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (23-Jul-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>On Jul 18, 2002 this sequence version replaced gi:20976361.</p>	<p>----- Genome Center</p> <p>Center: Baylor College of Medicine</p> <p>Center: BCM</p> <p>Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a></p> <p>Contact: hgsc-help@bcm.tmc.edu</p> <p>----- Project Information</p> <p>Center project name: GXM</p> <p>Center clone name: CH230-39G4</p> <p>----- Summary Statistics</p> <p>Sequencing vector: Plasmid;</p> <p>Chemistry: Dye-terminator Big Dye; 100% of reads</p> <p>Assembly program: Phrap; version 0.990329</p> <p>Consensus quality: 75985 bases at least Q40</p> <p>Consensus quality: 80102 bases at least Q30</p> <p>Consensus quality: 84340 bases at least Q20</p>	<p>-----</p> <p>* NOTE: Estimated insert size may differ from sequence length</p> <p>* (see <a href="http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html">http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html</a>)</p> <p>* NOTE: This is a 'working draft' sequence. It currently</p> <p>* consists of 52 contigs. The true order of the pieces</p> <p>* is not known and their order in this sequence record is</p> <p>* arbitrary. Gaps between the contigs are represented as</p> <p>* runs of 'N', but the exact sizes of the gaps are unknown.</p> <p>* This record will be updated with the finished sequence</p> <p>* as soon as it is available and the accession number will</p> <p>* be preserved.</p> <p>-----</p> <p>1 1157: contig of 1157 bp in length</p> <p>* 1158 1257: gap of unknown length</p> <p>* 1158 2612: contig of 1355 bp in length</p> <p>* 1213 2712: gap of unknown length</p> <p>* 2713 4192: contig of 1480 bp in length</p> <p>* 4193 4292: gap of unknown length</p> <p>* 4293 5374: contig of 1082 bp in length</p> <p>* 5375 5474: gap of unknown length</p> <p>* 5475 6615: contig of 1141 bp in length</p> <p>* 6616 6715: gap of unknown length</p> <p>* 6716 8340: contig of 1625 bp in length</p> <p>* 8341 8440: gap of unknown length</p> <p>* 8441 9488: contig of 1058 bp in length</p> <p>* 9489 9588: gap of unknown length</p> <p>* 9589 10938: contig of 1340 bp in length</p> <p>* 10939 11038: gap of unknown length</p> <p>* 11039 12186: contig of 1148 bp in length</p> <p>* 12187 12286: gap of unknown length</p> <p>* 12287 13914: contig of 1628 bp in length</p> <p>* 13915 14014: gap of unknown length</p> <p>* 14015 15174: contig of 1160 bp in length</p> <p>* 15175 15274: gap of unknown length</p> <p>* 15275 17059: contig of 1785 bp in length</p> <p>* 17060 17159: gap of unknown length</p> <p>* 17160 18388: contig of 1229 bp in length</p> <p>* 18389 18488: gap of unknown length</p> <p>* 18489 20246: contig of 1758 bp in length</p> <p>* 20247 20346: gap of unknown length</p> <p>* 20347 22074: contig of 1728 bp in length</p> <p>* 22075 22174: gap of unknown length</p> <p>* 22175 23874: contig of 1700 bp in length</p> <p>* 23875 23974: gap of unknown length</p> <p>* 23975 25681: contig of 1707 bp in length</p> <p>* 25682 27329: contig of 1548 bp in length</p> <p>* 27330 27429: gap of unknown length</p>

*	27430	29628:	contig of 2199 bp in length
*	29629	29728:	gap of unknown length
*	29729	31613:	contig of 1885 bp in length
*	31614	31713:	gap of unknown length
*	31714	33399:	contig of 1686 bp in length
*	33400	33499:	gap of unknown length
*	33500	33295:	contig of 2796 bp in length
*	36296	36395:	gap of unknown length
*	36396	37866:	contig of 1471 bp in length
*	37867	37966:	gap of unknown length
*	37967	39763:	contig of 1797 bp in length
*	39764	39863:	gap of unknown length
*	39864	41600:	contig of 1737 bp in length
*	41601	41700:	gap of unknown length
*	43251	43250:	contig of 1550 bp in length
*	43351	43350:	gap of unknown length
*	45414	45413:	contig of 2063 bp in length
*	45514	45513:	gap of unknown length
*	47724	47723:	contig of 2211 bp in length
*	47725	47824:	gap of unknown length
*	47825	50130:	contig of 2306 bp in length
*	50131	50230:	gap of unknown length
*	50231	52366:	contig of 2136 bp in length
*	52367	52466:	gap of unknown length
*	52467	54715:	contig of 2249 bp in length
*	54716	54815:	gap of unknown length
*	54816	57426:	contig of 2611 bp in length
*	57427	57526:	gap of unknown length
*	57527	59459:	contig of 1933 bp in length
*	59460	59559:	gap of unknown length
*	59560	62509:	contig of 2950 bp in length
*	62510	62609:	gap of unknown length
*	62610	64098:	contig of 1489 bp in length
*	64099	64198:	gap of unknown length
*	64199	66760:	contig of 2562 bp in length
*	66761	66860:	gap of unknown length
*	66861	69148:	contig of 2288 bp in length
*	69149	69248:	gap of unknown length
*	69249	70940:	contig of 1692 bp in length
*	70941	71040:	gap of unknown length
*	71041	73541:	contig of 2501 bp in length
*	73542	73641:	gap of unknown length
*	73642	75972:	contig of 2331 bp in length
*	75973	76072:	gap of unknown length
*	76073	78531:	contig of 2459 bp in length
*	78532	78631:	gap of unknown length
*	81304	81403:	contig of 2672 bp in length
*	81404	83837:	gap of unknown length
*	83838	83937:	gap of unknown length
*	83938	85968:	contig of 2031 bp in length
*	85969	86068:	gap of unknown length
*	86069	89935:	contig of 3867 bp in length
*	89936	90035:	gap of unknown length
*	90036	92970:	contig of 2934 bp in length
*	92970	93069:	gap of unknown length
*	93070	97516:	contig of 4447 bp in length
*	97517	97616:	gap of unknown length
*	97617	102656:	contig of 5040 bp in length
*	102657	102756:	gap of unknown length
*	102757	107151:	contig of 4395 bp in length
*	107152	107251:	gap of unknown length
*	107252	111024:	contig of 3773 bp in length
*	111025	111124:	gap of unknown length
*	111026	117076:	contig of 5952 bp in length
*	117077	117176:	gap of unknown length
*	117177	124282:	contig of 7106 bp in length.

Location/Qualifiers

Source

1. 124282  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"

Query Match

36.0%; Score 170.6; DB 2; Length 124282;

Best Local Similarity 81.7%; Pred. No. 6,8e-30;  
Matches 197; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy	31	GGATTAAATTTGGCCCATTTATCATCATAGTTTTTCTTATGAGACATGTTTATAGT	90
Db	33501	GGATTAACTGGTGCGCTTATCATCTTGTCTCTTAGAAGCATGTTTACAGT	32442
Qy	91	GTTCATCAAGTCCCTAACAGCACTGAATACGAAATCAGTTAAAAAGATGATC	150
Db	32441	GTTCATCAAGACACATTAACGACCGAAATPACAGAGAGTAAAGAGATGATC	32382
Qy	151	CTTGCCAAAGTTTTCTTTATAGTATTACGATTCGATTCGATGATCCATTTT	210
Db	32381	CTGCGCAAGCGTTTCTTTATGTTGTTTACGATGATTCGATTCATTTT	32322
Qy	211	GTTCGCAACCTTTTCTGCTTACGATGATTAATPACAGTACATTAACCTTTGGGTA	270
Db	32321	ATACGTAAATTTCTCTCAGTATGCGGTGAATPACAGTAAACCTTTCTGGGTA	32262
Qy	271	G 271	
Db	32261	G 32261	

RESULT 15  
AC098607  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-149B15, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 66 unordered pieces.  
ACCESSION  
AC098607.4 GI:21729834  
VERSION  
KEYWORDS  
HTG; HTGS PHASE1.  
SOURCE  
Norway rat.  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 173546)  
Munzy D.M., Adams C., Adio-Oduola B., Ali-oshman, F.R., Allen C.,  
Albrooke S.L., Amaralunga H.C., Are J.R., Ayale M., Banks T.,  
Barbataia J., Benton J., Bimege K., Blankenburg K., Bonnin D.,  
Bouck J., Bowle S., Brieva M., Brown E., Brown M., Bryant N.P.,  
Bunay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,  
Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,  
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,  
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,  
Dayila M.L., Davis C., Davy-Carroll L., Dederich D.A.,  
DeLaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,  
Doutwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,  
Eamhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,  
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Franz P.,  
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,  
Gottrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton R.,  
Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,  
Hernandez O., Hodgson A., Hognes M., Hollaway C., Hollins G.,  
Homesi F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E.,  
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,  
Karlsens E., Kelly S., Khan U., Kling L., Korvah J., Kovar C.,  
Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,  
Li J., Li Z., Licharge O., Lieu C., Liu D., Liu W., Loulesged H.,  
Lozato R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,  
Maheshwari M., Mapa P., Martin R., Matindale A., Martinez E.,  
Maser E., Mawhinney E., McLeod M.P., Meador M., Mei G., Metzger M.,  
Miner G., Miner Z., Mitchell T., Monabadi K., Morgan M., Morris S.,  
Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,  
Nguyen N., Nickerson E., Nwokwenko S., Ogih M., Okumou G.,  
Oregunye N., Oyiedo R., Pace A., Payton B., Peery J., Perez L.,  
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,  
Rivers M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Saverly G.,  
Scherer S., Scott G., Shen H., Shoshetitz N., Slason I.,  
Sodergren E., Sonakke T., Sparks A., Stanley H., Stone H.,  
Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,  
Tansey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,

Umanai, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, F., Zhou, T., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 173546)

Worley, K. C.

Direct Submission

Submitted (26-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 173546)

Worley, K. C.

Direct Submission

Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2002 this sequence version replaced gi:17973780.

----- Genome Center -----

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information -----

Center project name: GILC

Center clone name: CH230-148B15

----- Summary Statistics -----

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 115888 bases at least Q40

Consensus quality: 121372 bases at least Q30

Consensus quality: 125291 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_dir/alt\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_dir/alt_data.html))

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 66 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

*	1	1226:	contig of 1226 bp in length
*	1227	1226:	gap of unknown length
*	1327	2723:	contig of 1397 bp in length
*	2724	2883:	gap of unknown length
*	2824	4083:	contig of 1260 bp in length
*	4084	4183:	gap of unknown length
*	4184	5504:	contig of 1321 bp in length
*	5505	5604:	gap of unknown length
*	5605	6949:	contig of 1345 bp in length
*	6950	7049:	gap of unknown length
*	7050	8449:	contig of 1400 bp in length
*	8450	8549:	gap of unknown length
*	8550	10125:	contig of 1576 bp in length
*	10126	10225:	gap of unknown length
*	10226	11238:	contig of 1013 bp in length
*	11239	11338:	gap of unknown length
*	11339	12555:	contig of 1217 bp in length
*	12556	12655:	gap of unknown length
*	12656	13676:	contig of 1021 bp in length
*	13677	13776:	gap of unknown length
*	13777	14881:	contig of 1105 bp in length
*	14882	14981:	gap of unknown length
*	14982	16093:	contig of 1112 bp in length
*	16094	16193:	gap of unknown length
*	16194	18100:	contig of 1907 bp in length
*	18101	18200:	gap of unknown length
*	18201	19354:	contig of 1154 bp in length

*	19355	19454:	gap of unknown length
*	19455	20905:	contig of 1451 bp in length
*	20906	21005:	gap of unknown length
*	21006	22298:	contig of 1293 bp in length
*	22299	22398:	gap of unknown length
*	22399	23951:	contig of 1553 bp in length
*	22952	24051:	gap of unknown length
*	24052	25962:	contig of 1911 bp in length
*	25963	26062:	gap of unknown length
*	26063	27494:	contig of 1432 bp in length
*	27495	27594:	gap of unknown length
*	27595	29826:	contig of 2232 bp in length
*	29827	29926:	gap of unknown length
*	31440	31639:	contig of 1713 bp in length
*	31440	31799:	gap of unknown length
*	31740	33474:	contig of 1665 bp in length
*	33425	33524:	gap of unknown length
*	33525	34534:	contig of 1010 bp in length
*	34535	34634:	gap of unknown length
*	34635	36900:	contig of 2266 bp in length
*	36901	37000:	gap of unknown length
*	37001	39685:	contig of 2665 bp in length
*	39686	39785:	gap of unknown length
*	39786	42077:	contig of 2232 bp in length
*	42078	42177:	gap of unknown length
*	42178	43416:	contig of 1239 bp in length
*	43417	43516:	gap of unknown length
*	43517	45408:	contig of 1892 bp in length
*	45409	45508:	gap of unknown length
*	45509	48237:	contig of 2729 bp in length
*	48238	48337:	gap of unknown length
*	48339	50309:	contig of 1972 bp in length
*	50310	50409:	gap of unknown length
*	50410	52102:	contig of 1693 bp in length
*	52103	52202:	gap of unknown length
*	52203	54134:	contig of 1932 bp in length
*	54135	54234:	gap of unknown length
*	54235	56222:	contig of 1986 bp in length
*	56223	56322:	gap of unknown length
*	56323	58004:	contig of 1682 bp in length
*	58005	58104:	gap of unknown length
*	58105	59897:	contig of 1799 bp in length
*	59898	59997:	gap of unknown length
*	59999	61513:	contig of 1516 bp in length
*	61514	61613:	gap of unknown length
*	61614	63887:	contig of 2274 bp in length
*	63888	63987:	gap of unknown length
*	63989	66822:	contig of 2835 bp in length
*	66893	66922:	gap of unknown length
*	66923	69422:	contig of 2500 bp in length
*	69523	69523:	gap of unknown length
*	69523	72833:	contig of 3310 bp in length
*	72833	72933:	gap of unknown length
*	72933	75189:	contig of 2257 bp in length
*	75189	75289:	gap of unknown length
*	75290	79107:	contig of 3818 bp in length
*	79108	79207:	gap of unknown length
*	79208	81946:	contig of 2739 bp in length
*	81947	82046:	gap of unknown length
*	82047	84158:	contig of 2112 bp in length
*	84159	84259	gap of unknown length
*	84259	86521:	contig of 2263 bp in length
*	86522	86621:	gap of unknown length
*	86622	88042:	contig of 2421 bp in length
*	88043	89142:	gap of unknown length
*	89143	91306:	contig of 2164 bp in length
*	91307	91406:	gap of unknown length
*	91407	93952:	contig of 2546 bp in length
*	93953	94053:	gap of unknown length
*	94053	96967:	contig of 2915 bp in length
*	96968	97067:	gap of unknown length
*	97068	99738:	contig of 2671 bp in length
*	99739	99838:	gap of unknown length







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 3, 2003, 12:53:57 ; Search time 226 Seconds  
(without alignments)  
4723.218 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474  
Sequence: 1 gccacgattatcagtcgac.....ggagatgcacacgagttaa 474

Scoring table: IDENTITY NUC  
Gap 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues  
number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: N Geneset 101002.\*

```
1: /SID27/gcgdata/geneseq/geneq-emb1/NA1980.DAT.*
2: /SID27/gcgdata/geneseq/geneq-emb1/NA1981.DAT.*
3: /SID27/gcgdata/geneseq/geneq-emb1/NA1982.DAT.*
4: /SID27/gcgdata/geneseq/geneq-emb1/NA1983.DAT.*
5: /SID27/gcgdata/geneseq/geneq-emb1/NA1984.DAT.*
6: /SID27/gcgdata/geneseq/geneq-emb1/NA1985.DAT.*
7: /SID27/gcgdata/geneseq/geneq-emb1/NA1986.DAT.*
8: /SID27/gcgdata/geneseq/geneq-emb1/NA1987.DAT.*
9: /SID27/gcgdata/geneseq/geneq-emb1/NA1988.DAT.*
10: /SID27/gcgdata/geneseq/geneq-emb1/NA1989.DAT.*
11: /SID27/gcgdata/geneseq/geneq-emb1/NA1990.DAT.*
12: /SID27/gcgdata/geneseq/geneq-emb1/NA1991.DAT.*
13: /SID27/gcgdata/geneseq/geneq-emb1/NA1992.DAT.*
14: /SID27/gcgdata/geneseq/geneq-emb1/NA1993.DAT.*
15: /SID27/gcgdata/geneseq/geneq-emb1/NA1994.DAT.*
16: /SID27/gcgdata/geneseq/geneq-emb1/NA1995.DAT.*
17: /SID27/gcgdata/geneseq/geneq-emb1/NA1996.DAT.*
18: /SID27/gcgdata/geneseq/geneq-emb1/NA1997.DAT.*
19: /SID27/gcgdata/geneseq/geneq-emb1/NA1998.DAT.*
20: /SID27/gcgdata/geneseq/geneq-emb1/NA1999.DAT.*
21: /SID27/gcgdata/geneseq/geneq-emb1/NA2000.DAT.*
22: /SID27/gcgdata/geneseq/geneq-emb1/NA2001A.DAT.*
23: /SID27/gcgdata/geneseq/geneq-emb1/NA2001B.DAT.*
24: /SID27/gcgdata/geneseq/geneq-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474	100.0	474	22	AAAF28059
2	439.2	92.7	1089	22	AAI99557
3	439.2	92.7	1162	23	ABK3573
4	439.2	92.7	1191	22	AA06507
5	439.2	92.7	1804	21	AA200524
6	439.2	92.7	2467	20	AA225345
7	439.2	92.7	3584	20	AA225346
8	438.8	92.6	530	22	AAI99584
9	438.8	92.6	530	22	AAI35644

10	438.8	92.6	530	22	ABA06471	Human CDNA SEQ ID
11	438.8	92.6	530	22	AAS28950	CDNA encoding for
12	438.8	92.6	530	22	AAS29573	Human endocrine po
13	438.8	92.6	530	22	AAS30187	DNA encoding rena
14	438.8	92.6	530	22	AAS34845	CDNA encoding nove
15	438.8	92.6	530	22	ABK3875	DNA encoding novel
16	353.2	74.5	420	21	AACT5735	Human ORFX ORF1290
17	171.8	36.2	1065	22	ABL40194	Human G protein-co
18	171.8	36.2	1068	22	AAS07943	Human cDNA encodin
19	171.8	36.2	1473	24	ABL40192	Human G protein-co
20	171.8	36.2	1455	24	ABL40191	Human G protein-co
21	171.8	36.2	1830	24	ABL40188	Human G protein-co
22	171.8	36.2	2142	24	ABK51947	CDNA encoding huma
23	171.8	36.2	2190	24	ABL40197	Human G protein-co
24	171.8	36.2	2214	24	ABK51944	Human G protein-co
25	171.8	36.2	2262	24	ABL40196	CDNA encoding huma
26	156.8	33.1	321	24	AAAD32027	Human G protein-co
27	154.4	32.6	636	21	AAA44932	Human novel G-prot
28	106.8	22.5	1018	22	AAH51001	Human secreted exp
29	48.8	10.3	1015	22	AAS57085	Human ngpCR57 codi
30	48.8	10.3	1015	22	ABL07293	CDNA encoding Dros
31	44	9.3	4292	21	AAAF21792	Drosophila melanog
32	44	9.3	5644	24	ABK83490	Human breast and o
33	40.2	8.5	134499	21	AAAF22286	Human cDNA differe
34	39.8	8.4	113515	24	ABL34174	BAC containing rep
35	39.4	8.3	39536	23	ABL15482	Human immune syste
36	38.8	8.2	12138	24	ABK40034	Drosophila melanog
37	38.8	8.2	12138	24	ABL33629	Human chemically p
38	38.4	8.1	11797	20	AAH83965	Human immune syste
39	38.4	8.1	8967	20	AAH83938	Salmonella typhimu
40	38.4	8.1	8967	20	AAH83940	Salmonella typhimu
41	38.4	8.1	8967	20	AAH83942	Salmonella typhimu
42	38.4	8.1	8967	20	AAH83944	Salmonella typhimu
43	38.4	8.1	24701	20	AAH83935	Salmonella typhimu
44	38.4	8.1	24701	20	AAH83936	Salmonella typhimu
45	37.8	8.0	1205	18	AAH83358	Breast cancer tumo

#### ALIGNMENTS

RESULT 1	AAAF28059	standard; cDNA; 474 BP.
ID	AAAF28059;	
AC	AAAF28059;	
XX		
DT	23-MAY-2001	(first entry)
XX		
DE	Human HGRL101 G-protein coupled receptor coding sequence.	
XX		
KW	Human; HGRL101; G-protein coupled receptor; infection; pain; cancer;	
KW	diabetes; obesity; eating disorder; asthma; Parkinson's disease;	
KW	hypertension; osteoporosis; myocardial infarction; migraine; allergy;	
KW	psychotic disorder; neurological disorder; dyskinesia; vaccine; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..474
FT		/*tag= a
FT		/product= "HGRL101"
FT		/partial
XX		
PN	WO200114548-A2.	
XX		
PD	01-MAR-2001.	
XX		
PF	09-AUG-2000; 2000WO-EP07723.	
XX		
PR	19-AUG-1999; 99EP-0116345.	
XX		
PA	(MERE ) MERCK PATENT GMBH.	

```
XX
PI Duecker K;
XX
XX MPI: 2001-226617/23.
DR P-PSDB; AAB35407.
XX
XX Novel G-coupled protein receptor, HGR101 useful for treating diseases
PT such as microbial infections, cancers, obesity, asthma, diabetes,
PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer,
PT allergy -
XX
XX Claim 5; Page 35-36; 36pp; English.
XX
XX The present invention provides the protein and coding sequences for a
CC novel human G-protein coupled receptor, designated HGR101. The sequences
CC are useful in the diagnosis, prevention and treatment of diseases
CC including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
CC vomiting, psychotic and neurological disorders such as anxiety,
CC schizophrenia, manic depression, depression, delirium, dementia and
CC severe mental retardation, and dyskinesias including Huntington's disease
CC and Gilles de la Tourette's syndrome. The present sequence is the HGR101
CC partial coding sequence.
XX
XX Sequence 474 BP; 144 A; 92 C; 88 G; 150 T; 0 other;
SQ
Query Match 100.0%; Score 474; DB 22; Length 474;
Best Local Similarity 100.0%; Pred. No. 9,9e-119;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTTGGCCGATTTATCATCAT 60
Db 1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTTGGCCGATTTATCATCAT 60
QY 61 GTTTTTCCTATGAGCAAGTTTATAGTGTTCATCAAGTGCCATPAACAGACTGAA 120
Db 61 GTTTTTCCTATGAGCAAGTTTATAGTGTTCATCAAGTGCCATPAACAGACTGAA 120
QY 121 ATACGGAATCAAGTTAAAGAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 180
Db 121 ATACGGAATCAAGTTAAAGAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 180
QY 181 ACTGATGATTATGCTGATACCATTTTGTAGCGAAACCTCTTACGCTTCAGTGA 240
Db 181 ACTGATGATTATGCTGATACCATTTTGTAGCGAAACCTCTTACGCTTCAGTGA 240
QY 241 GAAATACAGGTACCATTAACCTCTGGGTAGTGAATGGTATTCGCCATTAAACAGTGT 300
Db 241 GAAATACAGGTACCATTAACCTCTGGGTAGTGAATGGTATTCGCCATTAAACAGTGT 300
QY 301 TTGAACCAATTCCTATACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTTT 360
Db 301 TTGAACCAATTCCTATACTCTGACCAAGACCAATTTAAAGAAATGATTCATCGGTTT 360
QY 361 TGGCATTAAGTACAGCAAGAAATCTATGAGACGCAAGATGATCAAGAAACATATGCTC 420
Db 361 TGGCATTAAGTACAGCAAGAAATCTATGAGACGCAAGATGATCAAGAAACATATGCTC 420
QY 421 CATCATTCATCTGGGGGAAATGTGGCCATGACAGAGATCCACTGAGTTAA 474
Db 421 CATCATTCATCTGGGGGAAATGTGGCCATGACAGAGATCCACTGAGTTAA 474
RESULT 2
AAI99557
ID AAI99557 standard; cDNA; 1089 BP.
XX
XX AAI99557;
AC
XX
XX 04-JAN-2002 (first entry)
XX
```

```
DE Human expressed polynucleotide SEQ ID NO 20.
XX
XX Human; nootropic; neuroprotective; cyrostatic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnetary;
KW antiparkinsonian; antislaking; antianaemic; antiarthritic; cancer;
KW antineumatic; hepatocytic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilester; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine;
KW 88.
XX
XX Homo sapiens.
XX
XX WO20015387-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01310.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218220.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0228509.
PR 05-SEP-2000; 2000US-0228511.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
```



Qy 361 TGGATACTACAGCAAGAAATCTATGACGCAAGGATACAGAAACATATGCTC 420  
Db 380 TGGTATACTACAGCAAGAAATCTATGACGCAAGG--TCAGAAACATATGCTC 437  
Qy 421 CATCATTCATCTGGGGAAATGTGCGACCTGACAGAGATGCCACTGATTTAA 474  
Db 438 CATCATTCATCTGGGTGGAAATGTGCGACCTGACAGAGATGCCACTGATTTAA 491  
RESULT 3  
ID ABK43573/c  
XX ABK43573 standard; cDNA; 1162 BP.  
AC ABK43573;  
XX  
XX 05-JUN-2002 (first entry)  
XX  
DE DNA encoding novel central nervous system protein #153.  
XX  
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
XX hyperproliferative disorder; neoplasm; cardiovascular disorder;  
XX cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
XX nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
XX acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
XX adenocarcinoma; reproductive system disorder; testicular feminisation;  
XX endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
XX respiratory disorder; renal disorder; kidney failure; blood disorder;  
XX myocardial infarction; wound healing; cell proliferation; skin aging;  
XX food additive; food preservative; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX W0200155318-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01332.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0230963.  
XX 26-JUL-2000; 2000US-0230964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0235758.  
XX 14-AUG-2000; 2000US-0235759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244611.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX  
 PI Rosen CA, Barash SC, Ruben SW;  
 XX  
 DR WPI: 2001-581633/65.  
 DR P-PSDB; AA087243.  
 XX  
 PT New isolated nucleic acid encoding a protein for diagnosing,  
 PT preventing, treating or ameliorating medical conditions and used as  
 PT food additives or preservatives -  
 XX  
 XX  
 PS Claim 1; SEQ ID No 163; 837bp; English.

CC The invention describes an isolated nucleic acid molecule (I) encoding a  
 CC novel central nervous system protein. (I) and polypeptides (II) encoded  
 CC by (I), are used to treat a medical conditions and in diagnosis of a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,  
 CC amyotrophic lateral sclerosis, nervous system disorders e.g. Alzheimer's disease and  
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
 CC adenocarcinomas and irritable bowel syndrome, reproductive system  
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
 CC acute kidney failure and blood related disorders e.g. myocardial  
 CC infarction. The polypeptides can also be used to aid wound healing and  
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
 CC maintain organs before transplantation, for supporting cell culture of  
 CC primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 92.7%; Score 439.2; DB 23; Length 1162;  
 Best Local Similarity 96.8%; Pred. No. 3.5e-109; Indels 2; Gaps 1;  
 Matches 459; Conservative 0; Mismatches 13;

QY 1 GCCGAGTTTATTCAGTGGCAATTTTCTTGATTAATTTGGCGCATTTATCATCATTA 60  
 DB 1080 GCCGAGTTTATTCAGTGGCAATTTTCTTGATTAATTTGGCGCATTTATCATCATTA 1021

QY 61 GTTTTTTCTTATGGAAGCATGTTTATAGTGTATCATCAAGTGCATPACAGCACTGAA 120  
 DB 1020 GTTTTTTCTTATGGAAGCATGTTTATAGTGTATCATCAAGTGCATPACAGCACTGAA 961  
 QY 121 ATTCGGAATCAAGTTTAAAAAGATGATCCTTGCACAAAGTTTTTCTTTATGTAATTT 180  
 DB 960 ATACGAATCAAGTTTAAAAAGATGATCCTTGCACAAAGTTTTTCTTTATGTAATTT 901  
 QY 181 ACTGATGATTTATGCTGTGATACCCATTTTGTAGGAAACCTTTTCACTGCTCAGGTA 240  
 DB 900 ACTGATGATTTATGCTGTGATACCCATTTTGTAGGAAATTTCTTCACTGCTCAGGTA 841  
 QY 241 GAAATPACAGGTACATPACCTCTTGGGTAGTATGTTGTTATTCGCCATTAACAGTCT 300  
 DB 840 GAAATPACAGGTACATPACCTCTTGGGTAGTATGTTTATTCGCCATTAACAGTCT 781  
 QY 301 TTGAACCAATTCCTATPACTGTGACCAAGACCATTTAAGAAATGATTCATCGGTTT 360  
 DB 780 TTGAACCAATTCCTATPACTGTGACCAAGACCATTTAAGAAATGATTCATCGGTTT 721  
 QY 361 TGGCATTACTACAGCAAAAGAAATCTATGACAGCAAGGATTCAGAAAATATGCTC 420  
 DB 720 TGGTATPACTACAGCAAAAGAAATCTATGACAGCAAGG--TCAGAAAATATGCTC 663  
 QY 421 CATCATTCATCTGGGGGAAATGTGGCCATGCGAGAGATGCCACGTGAGTTAA 474  
 DB 662 CATCATTCATCTGGGGGAAATGTGGCCATGCGAGAGATGCCACGTGAGTTAA 609

## RESULT 4

AA06507 standard; DNA; 1191 BP.

AA06507;

10-AUG-2001 (first entry)

Human COM222 G protein-coupled receptor DNA.

KW Human; G protein-coupled receptor; GPCR; COM222 protein; schizophrenia;  
 KW neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;  
 KW neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;  
 KW attention deficit hyperactivity disorder; neuropsychia; senile dementia;  
 KW affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;  
 KW depression; migraine; genetic screening; ds.

OS Homo sapiens.

Key Location/Qualifiers

CDS 1..1191

/\*tag= a /product= "Human G protein-coupled receptor protein"

MO200131014-A2.

03-MAY-2001.

27-OCT-2000; 2000MO-US29601.

PR 27-OCT-1999; 99US-0427653.  
 PR 27-OCT-1999; 99US-0427859.  
 PR 27-OCT-1999; 99US-0428020.  
 PR 27-OCT-1999; 99US-0428114.  
 PR 28-OCT-1999; 99US-0429517.  
 PR 28-OCT-1999; 99US-0429555.  
 PR 28-OCT-1999; 99US-0429676.  
 PR 28-OCT-1999; 99US-0429695.  
 PR 03-DEC-1999; 99US-0454399.  
 PR 12-JAN-2000; 2000US-0481794.

(PHAA) PHARMACIA & UPJOHN CO.

PI Vogeli G, Wood LS, Merchant K;

XX WPI, 2001-328653/34.

DR P-PSDB; AAB02498.

XX  
XX Seven transmembrane receptor polypeptides and polynucleotides, useful  
PT for treating neurological or psychiatric disorders, e.g. schizophrenia,  
PT as well as for identifying compounds useful for treating schizophrenia

PS Claim 22; Page 15-16; 215pp; English.

XX  
XX The invention relates to human G protein-coupled receptor (GPCR) and  
CC their corresponding DNA molecules. GPCR is also referred as seven  
CC transmembrane receptor. G protein-coupled receptor protein is useful for  
CC treating neurological disorder, particularly schizophrenia. GPCR protein  
CC is also useful for identifying compounds useful for treating  
CC schizophrenia. These compounds are also useful for treating other  
CC neurological and psychiatric diseases, e.g. depression, anxiety, bipolar  
CC disease, affective disorders, attention deficit hyperactivity disorder/  
CC attention deficit disorder, epilepsy, neuritis, neuroasthenia, neuropathy,  
CC neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile  
CC dementia. The invention also provides genetic screening procedures that  
CC entail analysing a person's genome with respect to GPCR. The vectors are  
CC useful for the recombinant production of the GPCR's. The present DNA  
CC sequence encodes human CON222 G protein-coupled receptor (GPCR) protein.

XX Sequence 1191 BP; 340 A; 229 C; 226 G; 396 T; 0 other;

Query Match 92.7%; Score 439.2; DB 22; Length 1191;

Best Local Similarity 96.8%; Pred. No. 3.5e-109;

Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTTCTGTATTAATTGGCCGATTTATCATATA 60  
DB 643 GCCCAGATTATTCAGTGGCAATTTTCTGTATTAATTGGCCGATTTATCATATA 702  
QY 61 GTTTTTCCTATGAGAGATGTTTATAGTGTATCATCAAGTGCATACAGCACTGAA 120  
DB 703 GTTTTTCCTATGAGAGATGTTTATAGTGTATCATCAAGTGCATACAGCACTGAA 762  
QY 121 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAAGCTTTTCTTATATATTT 180  
DB 763 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAAGCTTTTCTTATATATTT 822  
QY 181 ACTGATGATTAATGCTGAGTACCATTTTGTAGGAAATTTCTTATCTGCTTCAAGTA 240  
DB 823 ACTGATGATTAATGCTGAGTACCATTTTGTAGGAAATTTCTTATCTGCTTCAAGTA 882  
241 GAAATACCAAGTACATTAACCTTGGTATGATGTTATCTGCAATTAACAGTGCT 300  
DB 883 GAAATACCAAGTACATTAACCTTGGTATGATGTTATCTGCAATTAACAGTGCT 942  
QY 301 TTGAACCAATCTCTATATCTGTACACAGACCAATTAAGAAATGATTCATGGTTT 360  
DB 943 TTGAACCAATCTCTATATCTGTACACAGACCAATTAAGAAATGATTCATGGTTT 1002  
QY 361 TGGCATTAATCAAGCAAGAAATCTATGACAGCAAGGATATCAAGAAATATGCTTC 420  
DB 1003 TGGTATTAATCAAGCAAGAAATCTATGACAGCAAGG-TCAGAAATCATATGCTTC 1060  
QY 421 CATCATTAATCTGGGGGAATGTGGCCATGCAAGAGATGCCCTGAGTTAA 474  
DB 1061 CATCATTAATCTGGGGGAATGTGGCCATGCAAGAGATGCCCTGAGTTAA 1114

RESULT 5

AA290524

ID AA290524 standard; cDNA; 1804 BP.

XX AA290524;

XX 05-JUN-2000 (first entry)

XX  
DE Human GPCR protein (HGPRP) encoding cDNA (clone ID 2488822).

XX  
XX Human; G protein coupled protein receptor. HGPRP; cell proliferation;  
XX neurological; immune disorder; cytostatic; anti-arteriosclerotic;  
XX anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;  
XX immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;  
XX anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.

OS Homo sapiens.

PN WO200015793-A2.

XX 23-MAR-2000.

XX 17-SEP-1999; 99MO-US20958.

XX 17-SEP-1998; 98US-0156513.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

PI Baughn MR;

DR WPI; 2000-271432/23.

P-PSDB; AAY57286.

PT Human G protein coupled protein receptor peptides useful for the  
PT prevention, diagnosis and treatment of cell proliferative, neurological  
PT and immune disorders -

PS Claim 9; Page 69-70; 71pp; English.

XX  
XX The invention provides human G protein coupled protein receptor (HGPRP)  
CC polypeptides and polynucleotides encoding them. The polypeptides can be  
CC produced by standard recombinant methodology. The polynucleotides and  
CC polypeptides may be used in the prevention, treatment and diagnosis of  
CC diseases associated with their inappropriate expression. Diseases that  
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,  
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and  
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's  
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease  
CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as  
CC diagnostic agents for detecting the presence of HGPRP polypeptides in  
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences  
CC AA290521-526 represent cDNA fragments encoding the HGPRP polypeptides  
CC (AAY57283-288).

SQ Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;

Query Match 92.7%; Score 439.2; DB 21; Length 1804;

Best Local Similarity 96.8%; Pred. No. 4e-109;

Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTTCTGTATTAATTGGCCGATTTATCATATA 60  
DB 792 GCCCAGATTATTCAGTGGCAATTTTCTGTATTAATTGGCCGATTTATCATATA 851  
QY 61 GTTTTTCCTATGAGAGATGTTTATAGTGTATCATCAAGTGCATACAGCACTGAA 120  
DB 852 GTTTTTCCTATGAGAGATGTTTATAGTGTATCATCAAGTGCATACAGCACTGAA 911  
QY 121 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAAGCTTTTCTTATATATTT 180  
DB 912 ATACGGAATCAAGTTAAAGAGATGATCCTTGGCAAGCTTTTCTTATATATTT 971  
QY 181 ACTGATGATTAATGCTGAGTACCATTTTGTAGGAAATTTCTTATCTGCTTCAAGTA 240  
DB 972 ACTGATGATTAATGCTGAGTACCATTTTGTAGGAAATTTCTTATCTGCTTCAAGTA 1031  
QY 241 GAAATACCAAGTACATTAACCTTGGTATGATGTTATCTGCAATTAACAGTGCT 300  
DB 1032 GAAATACCAAGTACATTAACCTTGGTATGATGTTATTTATCTGCAATTAACAGTGCT 1091



QY 301 TTGAACCAATTCCTATATCTGACCAAGACCATTTAAAGAAATGATTCATGGTTT 360  
DB 1092 TTGAACCAATTCCTATATCTGACCAAGACCATTTAAAGAAATGATTCATGGTTT 1151  
QY 361 TGGCATTAATCTACAGACAAAGAAATCTATGACAGCAAGATATCAGAAACATATGCTC 420  
DB 1152 TGGTATTAATCTACAGACAAAGAAATCTATGACAGCAAGG--TCGAAAACATATGCTC 1209  
QY 421 CATCATTCATCTGGGGGGAATGTGGCCACTGAGAGATGTCACCTGATTAA 474  
DB 1210 CATCATTCATCTGGGGGGAATGTGGCCACTGAGAGATGTCACCTGATTAA 1263

RESULT 6  
AAZ25345  
ID AAZ25345 standard; cDNA; 2467 BP.  
AC AAZ25345;  
AC 20-DEC-1999 (first entry)

DE Human LGR7 long form nucleotide sequence.  
XX  
XX Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;  
KM extracellular leucine rich repeat region; mapping; identification; ss.  
XX  
XX Homo sapiens.  
XX  
XX MO9948921-A1.  
XX  
XX 30-SEP-1999.  
XX  
XX 25-MAR-1999; 99WO-US06573.  
XX  
XX 26-MAR-1998; 98US-0079501.  
XX  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
PA (ORGA ) ORGANON NV.  
XX  
XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;  
PI WPI; 1999-591074/50.  
XX  
XX P-PSDB; AAY42170.  
XX  
XX New G-protein coupled receptors, useful for identifying their own  
PT ligands -  
XX

Claim 4; Fig 3; 54pp; English.

CC The present sequence encodes the human G-protein coupled receptor  
CC having extracellular leucine rich repeat regions, designated LGR7 long  
CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for  
CC the receptor. The polypeptides and/or polynucleotides are also useful  
CC for homologous or related genes, producing compositions that modulate  
CC the expression or function of the receptors, gene therapy, mapping  
CC functional regions of the receptors, studying associated physiological  
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens  
CC for producing antibodies, and for identifying biologically active  
CC agents. The polypeptides contain a G-protein coupled seven  
CC transmembrane region and a leucine rich repeat extracellular domain.  
CC These regions capture and facilitate optimal orientation of its ligand.  
CC The proteins are also expressed in diverse tissues.  
XX  
XX  
XX Sequence 2467 BP; 747 A; 487 C; 474 G; 759 T; 0 other;  
SQ

Query Match 92.7%; Score 439.2; DB 20; Length 2467;  
Best Local Similarity 96.8%; Pred. No 4.4e-109;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCGAGATTATTCAGTGGCAATTTCTTGATTAATTTGGCCGATTTATCATCATATA 60  
DB 1919 GCCGAGATTATTCAGTGGCAATTTCTTGATTAATTTGGCCGATTTATCATCATATA 1978

QY 61 GTTTTTCCTATGAGACATGTTTATATGTTTCATCAAAAGTCCATACAGCAACTGAA 120  
DB 1979 GTTTTTCCTATGAGACATGTTTATATGTTTCATCAAAAGTCCATACAGCAACTGAA 2038  
QY 121 ATACGAATCAAGTTAAAGAAAGATGATCTTGGCCAAAGTTTCTTTATATATTTT 180  
DB 2039 ATACGAATCAAGTTAAAGAAAGATGATCTTGGCCAAAGTTTCTTTATATATTTT 2098  
QY 181 ACTGATCATTAATGCTGATATACCATTTTGTAGCGAAACCTCTTCACTGCTTCAAGTA 240  
DB 2099 ACTGATCATTAATGCTGATATACCATTTTGTAGCGAAATTTCTTCACTGCTTCAAGTA 2158  
QY 241 GAAATACAGATACATTAACCTCTTGGGTATGATGTTTATTCATTAACAGTGCT 300  
DB 2159 GAAATACAGATACATTAACCTCTTGGGTATGATGTTTATTCATTAACAGTGCT 2218  
QY 301 TTGAACCAATTCCTATATCTGACCAAGACCATTTAAAGAAATGATTCATGGTTT 360  
DB 2219 TTGAACCAATTCCTATATCTGACCAAGACCATTTAAAGAAATGATTCATGGTTT 2278  
QY 361 TGGCATTAATCTACAGACAAAGAAATCTATGACAGCAAGATATCAGAAACATATGCTC 420  
DB 2279 TGGTATTAATCTACAGACAAAGAAATCTATGACAGCAAGG--TCGAAAACATATGCTC 2336  
QY 421 CATCATTCATCTGGGGGGAATGTGGCCACTGAGAGATGTCACCTGATTAA 474  
DB 2337 CATCATTCATCTGGGGGGAATGTGGCCACTGAGAGATGTCACCTGATTAA 2390

RESULT 7  
AAZ25346  
ID AAZ25346 standard; cDNA; 3584 BP.  
AC AAZ25346;  
AC 20-DEC-1999 (first entry)

DE Human LGR7 short form nucleotide sequence.  
XX  
XX Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;  
KM extracellular leucine rich repeat region; mapping; identification; ss.  
XX  
XX Homo sapiens.  
XX  
XX MO9948921-A1.  
XX  
XX 30-SEP-1999.  
XX  
XX 25-MAR-1999; 99WO-US06573.  
XX  
XX 26-MAR-1998; 98US-0079501.  
XX  
XX (STRD ) UNIV LELAND STANFORD JUNIOR.  
PA (ORGA ) ORGANON NV.  
XX  
XX Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;  
PI WPI; 1999-591074/50.  
XX  
XX P-PSDB; AAY42171.  
XX  
XX New G-protein coupled receptors, useful for identifying their own  
PT ligands -  
XX  
XX  
XX Claim 4; Fig 4; 54pp; English.

CC The present sequence encodes the human G-protein coupled receptor  
CC having extracellular leucine rich repeat regions, designated LGR7 short  
CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for  
CC the receptor. The polypeptides and/or polynucleotides are also useful  
CC for homologous or related genes, producing compositions that modulate  
CC the expression or function of the receptors, gene therapy, mapping  
CC functional regions of the receptors, studying associated physiological

CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens  
CC for producing antibodies, and for identifying biologically active  
CC agents. The polypeptides contain a G-protein coupled seven  
CC transmembrane region and a leucine rich repeat extracellular domain.  
CC These regions capture and facilitate optimal orientation of its ligand.  
CC The proteins are also expressed in diverse tissues.

XX Sequence 3584 BP; 1124 A; 670 C; 647 G; 1139 T; 4 other;

Query Match 92.7%; Score 439.2; DB 20; Length 3584;  
Best Local Similarity 96.8%; Pred. No. 4,9e-109;  
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTCTTGTAATTAATTGGCCCATTTATCATCA 60  
DB 1744 GCCCAGATTATTCAGTGGCAATTTCTTGTAATTAATTGGCCCATTTATCATCA 1803

QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGGCCATTAACAGCACTGAA 120  
DB 1804 GTTTTTCCTATGGAAGCATGTTTATAGTTCATCAAGGCCATTAACAGCACTGAA 1863

QY 121 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGTTTCTTATATATTT 180  
DB 1864 ATACGGAATCAAGTAAAGAGATGATCCTTGCCAAAGTTTCTTATATATTT 1923

QY 181 ACTGATGATTAAGTGGATACCCATTTTGTAGCCAAACCTCTTCAGCTTCAGGTA 240  
DB 1924 ACTGATGATTAAGTGGATACCCATTTTGTAGCCAAACCTCTTCAGCTTCAGGTA 1983

QY 241 GAAATACAGATACCATACCTCTGGGTAGTATGTTATTCGACATTAACAGTGCT 300  
DB 1984 GAAATACAGATACCATACCTCTGGGTAGTATGTTATTCGACATTAACAGTGCT 2043

QY 301 TTGAACCAATTCCTATCTGACCAAGACATTTAAAGAAATGATTCGGTTT 360  
DB 2044 TTGAACCAATTCCTATCTGACCAAGACATTTAAAGAAATGATTCGGTTT 2103

QY 361 TGGCATTAATCAAGCAAGAAATCTATGACGCAAGGATACGAAATATGATGCTC 420  
DB 2104 TGGCATTAATCAAGCAAGAAATCTATGACGCAAGGATACGAAATATGATGCTC 2161

QY 421 CATGATTCATGCGGGGAATGTGGCACTGACAGAGATGCCACTGAGTTAA 474  
DB 2162 CATGATTCATGCGGGGAATGTGGCACTGACAGAGATGCCACTGAGTTAA 2215

## RESULT 8

AAI99584 standard; cDNA; 530 BP.

AAI99584;

04-JAN-2002 (first entry)

Human expressed polymucleotide SEQ ID NO 47.

XX Human; nootropic; neuroprotective; cytoskeletal; dermatological; virocidic;  
XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnerary;  
XX antiparkinsonian; antitickling; antianemic; antiarthritic; cancer;  
XX antineurotic; hepatotropic; cerebroprotective; antiinflammatory;  
XX antileukemic; antidiabetic; antitumor; anticonvulsant; antifungal;  
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; nephrotropic; gene therapy; vaccine;  
XX 85.

XX Homo sapiens.

XX WO200155387-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01310.

PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184564.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218280.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226281.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233403.  
PR 14-SEP-2000; 2000US-0233404.  
PR 14-SEP-2000; 2000US-0233405.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234977.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-024121.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259676.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM,  
XX  
DR WPI: 2001-465573/50.  
XX P-PSDB; AAM99972.  
XX  
PT Isolated digestive system associated polypeptide for treating,  
preventing and/ or prognosing disorders related to the digestive system

PT including digestive system cancers and also for testing and detection  
PT e.g. diagnosis -  
XX  
XX Claim 1; SEQ ID NO 47; 509pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (AA199548-AA199604) and proteins  
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC hemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 92.6%; Score 438.8; DB 22; Length 530;  
Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
Matches 458; Conservative 1; Mismatches 11; Indels 2; Gaps 1;

QY 1 GCCGAGTTTATTCAGTGGCAATTTTCTGTGATTAATTTGGCCGATTTATCATCA 60  
DB 2 GCCGAGTTTATTCAGTGGCAATTTTCTGTGATTAATTTGGCCGATTTATCATCA 61  
QY 61 GTTTTTCCTTGAAGCAGTTTAAAGTGTTCATCAAGTCCATAACAGACTGAA 120  
DB 62 GTTTTTCCTTGAAGCAGTTTAAAGTGTTCATCAAGTCCATAACAGACTGAA 121  
QY 121 ATACGAGTCAAGTTAAAGAGATGATCCTTCCAAAGTTTCTTATATGATTT 180  
DB 122 ATACGAGTCAAGTTAAAGAGATGATCCTTCCAAAGTTTCTTATATGATTT 181  
QY 181 ACTGATGATTAATGCTGATTAATCCATTTTGTACGAAACCTTTTCACTGCTTCAGTA 240  
DB 182 ACTGATGATTAATGCTGATTAATCCATTTTGTGATGAAATTTTTCACGCTTCAGTA 241  
QY 241 GAAATACCAAGTACCATTAATCTTGGTGTGATTTGATTTCTGCAATTAACAGTGT 300  
DB 242 GAAATACCAAGTACCATTAATCTTGGTGTGATTTGATTTCTGCAATTAACAGTGT 301  
QY 301 TTGAACCAATTTCTATATCTGATCAACCAAGACCATTTAAAGAAATGATTCATCGGTT 360  
DB 302 TTGAACCAATTTCTATATCTGATCAACCAAGACCATTTAAAGAAATGATTCATCGGTT 361  
QY 361 TGGCATTAATCAAGCAAGAAATCTATGACGCAAGAGTATCGAAATCATATGCTC 420  
DB 362 TGGTATTAATCAAGCAAGAAATCTATGACGCAAGAG--TCAGAAATCATATGCTC 419  
QY 421 CATCATTAATCTGGGGGAAATGTGACCATGACGAGATGCCACTGAGTTAA 474  
DB 420 CATCATTAATCTGGGGGAAATGTGACCATGACGAGATGCCACTGAGTTAA 473  
RESULT 9  
AAL35644  
ID AAL35644 strand; cDNA; 530 BP.  
XX  
XX AAL35644;  
AC  
XX  
XX 08-JAN-2002 (first entry)  
DT  
XX  
XX Human musculoskeletal system related polynucleotide SEQ ID NO 986.  
DE  
XX

KM	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW	antiallergic; hepatotropic; antidiabetic; antineoplastic; antitumor;
KW	vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW	cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder
KW	neurological disease; infection; human; secreted protein;
KX	musculoskeletal system; ss.
XX	
OS	Homo sapiens.
PN	WO200155367-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01338.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	18-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226681.
PR	22-AUG-2000; 2000US-0226682.
PR	22-AUG-2000; 2000US-0227188.
PR	23-AUG-2000; 2000US-0227189.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.
PR	08-SEP-2000; 2000US-0231242.
PR	08-SEP-2000; 2000US-0231243.
PR	08-SEP-2000; 2000US-0231244.
PR	08-SEP-2000; 2000US-0231413.
PR	08-SEP-2000; 2000US-0231414.
PR	08-SEP-2000; 2000US-0232080.
PR	12-SEP-2000; 2000US-0232081.
PR	14-SEP-2000; 2000US-0232397.
PR	14-SEP-2000; 2000US-0232398.
PR	14-SEP-2000; 2000US-0232399.
PR	14-SEP-2000; 2000US-0232400.
PR	14-SEP-2000; 2000US-0232401.
PR	14-SEP-2000; 2000US-0233063.

[illegible]

PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-451937/48.  
DR P-PSDB; ABB04062.  
XX  
PT Isolated polypeptide for treating, preventing and/or diagnosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
PS Claim 1; SEQ ID NO 986; 781bp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune hemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPo at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;  
XX  
Query Match 92.6%; Score 438.8; DB 22; Length 530;  
Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
XX  
QY 1 GCCCAGATTATTCAGGCGCAATTTTCTGGTAATATTTGGCCGATTATTCATCATA 60  
2 GCCCAGATTATTCAGGCGCAATTTTCTGGTAATATTTGGCCGATTATTCATCATA 61  
QY 61 GTTTTTCCTATGAGAGCATGTTTATAGTTCATCAAGTCCATTAACAGCACTGAA 120  
DB 62 GTTTTTCCTATGAGAGCATGTTTATAGTTCATCAAGTCCATTAACAGCACTGAA 121  
QY 121 ATACGGAATCAAGTAAAAAGAGATATCTTGGCCAAAGTTTTCCTTAATAGTATT 180  
DB 122 ATACGGAATCAAGTAAAAAGAGATATCTTGGCCAAAGTTTTCCTTAATAGTATT 181  
QY 181 ACTGATGATTAATGCTGATATCCATTTTGTAGCGAAACCTCTTCAAGCTTCAGGTA 240  
DB 182 ACTGATGATTAATGCTGATATCCATTTTGTAGCGAAATTTCTTCAAGCTTCAGGTA 241  
QY 241 GAAATACAGAGTACCATTAACCTTTGGGTAGTATTGTTATTCGCATTAACAGTGCT 300  
DB 242 GAAATACAGAGTACCATTAACCTTTGGGTAGTATTGTTATTCGCATTAACAGTGCT 301  
QY 301 TTGAACCATTTCTCTATATCTCGACCAAGACCATTTTAAAGAAATGATTCGGTTT 360  
DB 302 TTGAACCATTTCTCTATATCTCGACCAAGACCATTTTAAAGAAATGATTCGGTTT 361  
QY 361 TGGCATTAATCAAGACAAAGAAATCTATGACAGCAAGATGATGAGAAACATATGCTC 420  
DB 362 TGGTATTAATCAAGACAAAGAAATCTATGACAGCAAGG--TCAGAAACATATGCTC 419

QY 421 CATCATTCATCTGGGGGGAAGTGGCCACTGACAGAGATGCCACTGAGTTAA 474  
DB 420 CATCATTCATCTGGGGGGAAGTGGCCACTGACAGAGATGCCACTGAGTTAA 473  
RESULT 10  
ABA06471  
ID ABA06471 standard; cDNA, 530 BP.  
XX  
AC ABA06471;  
XX  
XX 10-JAN-2002 (first entry)  
DE Human cDNA SEQ ID NO: 137.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation; SS.  
XX  
OS Homo sapiens.  
XX  
PN WO200154474-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01349.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 16-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190076P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214866P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216800P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225477P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 14-AUG-2000; 2000US-225759P.  
PR 18-AUG-2000; 2000US-226279P.  
PR 22-AUG-2000; 2000US-226681P.  
PR 22-AUG-2000; 2000US-226682P.  
PR 22-AUG-2000; 2000US-227182P.  
PR 23-AUG-2000; 2000US-227092P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-228927P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 06-SEP-2000; 2000US-230437P.  
PR 06-SEP-2000; 2000US-230438P.



Oy 421 CATCATTCATCTGGGGGAAATGTGGCCTGCAGAGATGCCACTGACTTAA 474  
Db 420 CATCATTCATCTGGGGTGGAAATGTGGCCACTGCAGAGATGCCACTGACTTAA 473

## RESULT 11

AA528950  
ID AA528950 standard; cDNA; 530 BP.

XX AA528950;

XX 21-NOV-2001 (first entry)

XX cDNA encoding for human uterine motility-association polypeptide #15.

KW Human; uterine motility-association disorder; uterus; pregnancy;

XX labour; menstrual cycle; gene therapy; ss.

XX Homo sapiens.

OS WO200155201-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01317.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0228287.

PR 01-SEP-2000; 2000US-0228343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
DR MPI; 2001-488777/53.  
DR P-PSDB; AAU18108.  
XX  
XX Isolated polypeptide and nucleic acid molecules for treating,  
PT preventing and/or prognosing disorders related to uterine motility  
XX e.g. disorders associated with pregnancy and the menstrual cycle -  
XX  
XX Claim 4; SEQ ID No 25; 524pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
CC uterine motility-association polypeptides (AAU18094-AAU18152),  
CC and cDNA and genomic sequences encoding for these polypeptides.  
CC The sequences of the invention are useful in the diagnosis,  
CC treatment, prevention and/or prognosis of diseases associated  
CC with uterine motility such as pregnancy and labour, and menstrual  
CC disorders. The polynucleotide sequences of the invention are also  
CC useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences  
CC encoding for novel human uterine motility-association polypeptides.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;  
SQ  
Query Match 92.6%; Score 438.8; DB 22; Length 530;  
Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
XX  
1 GCCCAGATTATTCAGGCGCAATTTCTTGATTAATTTGGCCGATTATCATCATCA  
DB 2 GCCCAGATTATTCAGGCGCAATTTCTTGATTAATTTGGCCGATTATCATCATCA 61  
XX  
61 GTTTTTCCTATGAGAGCATTTTATAGTTCATCAAAAGTGCATTAACGAACTGAA 120  
DB 62 GTTTTTCCTATGAGAGCATTTTATAGTTCATCAAAAGTGCATTAACGAACTGAA 121  
XX  
121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCTCAAGCTTTTCTTATAGTATT 180  
DB 122 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCTCAAGCTTTTCTTATAGTATT 181  
XX  
181 ACTGATGATTAATCTGTGATACCAATTTTGTAGGAAATTTCTTACAGCTTCAGGTA 240  
DB 182 ACTGATGATTAATCTGTGATACCAATTTTGTAGGAAATTTCTTACAGCTTCAGGTA 241  
XX  
241 GAAATACCAAGTATACCTCTTGAGTAGTATGTTATTTGTCATTAAACAGTCT 300  
DB 242 GAAATACCAAGTATACCTCTTGAGTAGTATGTTATTTGTCATTAAACAGTCT 301  
XX  
301 TTGAACCAATTTCTATAGTCTGACCAAGACATTTAAAGAAATGATTCATCGGTTT 360  
DB 302 TTGAACCAATTTCTATAGTCTGACCAAGACATTTAAAGAAATGATTCATCGGTTT 361

QY 361 TGGCATACTACAGACAAAGAAATCTATGACAGCAAGGTATCGAAGAAACATATGCTC 420  
DB 362 TGGTATATACAGACAAAGAAATCTATGACAGCAAGG--TCAGAAACATATGCTC 419  
QY 421 CATCATTTTCATCTGGGGGAAATGTGGCCACTGAGAGATGCGACCTGAGTTAA 474  
DB 420 CATCATTTTCATCTGGGGGAAATGTGGCCACTGAGAGATGCGACCTGAGTTAA 473  
RESULT 12  
AAS29573 standard; cDNA, 530 BP.  
ID AAS29573;  
XX  
XX AAS29573;  
AC  
XX  
DT 21-NOV-2001 (first entry)  
XX  
XX Human endocrine polypeptide encoding cDNA SEQ ID No 73.  
XX  
XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive;  
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antineumatic; antiproliferative; cytostatic; cardiac; neuroprotective;  
KW cerebroprotective; noctropic; antibacterial; fungicide; cancer;  
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; nervous system disorder; bacterial infection;  
KW cerebrovascular disorder; viral infection; ocular disorder; endocrine disorder;  
KW fungal infection; renal disorder; respiratory disorder;  
KW gastrointestinal disorder; organ transplantation; food preservative;  
KW wound healing; skin aging; anti-infertility;  
KW tissue regeneration; anti-infertility.  
XX  
XX Homo sapiens.  
OS  
XX MO200155364-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US01308.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0224521.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.



PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0232081.  
 PR 14-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232337.  
 PR 14-SEP-2000; 2000US-0232338.  
 PR 14-SEP-2000; 2000US-0232339.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235835.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 01-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249246.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251866.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0255678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-451936/48.  
 DR P-PSDB; AAU18344.  
 XX  
 PT Isolated polypeptide for treating, preventing and/or prognosing  
 PT disorders of the endocrine system such as reproductive disorders,  
 PT endocrine cancers and also for testing and detection e.g. diagnosis -  
 XX  
 PS Claim 1; SEQ ID No 73; 604pp; English.  
 XX  
 CC Sequences AAS29511-AAS29736 represent cDNA molecules, which encode the  
 CC endocrine polypeptides of the invention. Endocrine polypeptides and their  
 CC associated polynucleotides of the invention are useful in the diagnosis,  
 CC treatment and prevention of various types of disorders in e.g. humans,  
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A  
 CC pathological condition can be determined by determining the presence or  
 CC absence of a mutation in an endocrine polynucleotide. The treatable  
 CC disorders include autoimmune diseases such as rheumatoid arthritis,  
 CC hyperproliferative disorders such as neoplasms of the breast or liver,  
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular  
 CC disorders such as cerebral ischaemia, nervous system disorders such as  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,  
 CC ocular disorders such as corneal infection, endocrine disorders such as  
 CC premature labour and infertility, gastrointestinal disorders such as  
 CC Crohn's disease, renal disorders such as glomerulonephritis and  
 CC respiratory disorders such as asthma. The polypeptides can also be used  
 CC to aid wound healing, to prevent skin aging due to sunburn, to maintain  
 CC organs before transplantation, to regenerate tissues and in chemotaxis.  
 CC The polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO

Query Match 92.6%; Score 438.8; DB 22; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTTGGCCGCAATTATCATATA 60  
 Db 2 GCCCAGATTATTCAGTGGCAATTTTCTTGATTAATTTGGCCGCAATTATCATATA 61

QY 61 GTTTTTCCTATGGAGCATGTTTATAGTTCATCAAGTGCATPAACAGACACTGAA 120  
Db 62 GTTTTTCCATATGGAGCAGATTTTATAGTTCATCAAGTGCATPAACAGACACTGAA 121  
QY 121 ATACGAATCAAGTTAAAGAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 180  
Db 122 ATACGAATCAAGTTAAAGAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 181  
QY 181 ACTGATGCATTTAGCTGGATACCATTGTTGAGCGAAACCTCTTTCAGCTTCAGGTA 240  
Db 182 ACTGATGCATTTAGCTGGATACCATTGTTGAGCGAAACCTCTTTCAGCTTCAGGTA 241  
QY 241 GAATATACAGGTATACCATACCTCTGGGTAGTGTGATTTGATTCCTGCAATTAACAGTCT 300  
Db 242 GAATATACAGGTATACCATACCTCTGGGTAGTGTGATTTGATTCCTGCAATTAACAGTCT 301  
QY 301 TTGAACCCAAATTCCTATCTCTGACCAAGACCATTTTAAAGAAATGATTCACTGGTTT 360  
Db 302 TTGAACCCAAATTCCTATCTCTGACCAAGACCATTTTAAAGAAATGATTCACTGGTTT 361  
QY 361 TGGCATAACTACAGACAAAGAAATCTATGACACGCAAGGTATCAGAAACATATGCTC 420  
Db 362 TGGTATTAATCTACAGACAAAGAAATCTATGACACGCAAGGTATCAGAAACATATGCTC 419  
QY 421 CATGATTCATCTGGGGGAAATGCGCACCTGACAGAGATGCCACCTGAGTTAA 474  
Db 420 CATGATTCATCTGGGGGAAATGCGCACCTGACAGAGATGCCACCTGAGTTAA 473  
RESULT 13  
AAS30187  
ID AAS30187 standard; cDNA; 530 BP.  
AC AAS30187;  
XX 21-NOV-2001 (first entry)  
DT 21-NOV-2001 (first entry)  
XX  
DE DNA encoding renal and cardiovascular-associated protein, Seq ID 33.  
XX  
KW Human; antiinflammatory; neuroprotective; immunomodulator; vulnery;  
KW cardiovascular; cytosolic; nephrotropic; antianemic; nephritis;  
KW immunosuppressive; kidney disorder; renal failure; hypertension;  
KW cardiovascular disorder; myocardial infarction; blood disorder; anaemia;  
KW blood coagulation disorder; electrolyte imbalance disorder; cancer;  
KW hypoaemia; hyperkalaemia; neoplastic disease; reproductive system disorder;  
KW autoimmune disease; inflammatory disease; endocrine disorder;  
KW endocrine disorder; neural activity; neurological disorder;  
wound healing; respiratory disorder; ss.  
OS Homo sapiens.  
XX  
PN W020015328-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01359.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-018664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214866.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227189.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249246.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-488787/53.  
 DR P-PSDB; AAU18666.  
 XX  
 PT New polynucleotides and polypeptides, useful for diagnosing, treating,  
 PT preventing or prognosing e.g. kidney, cardiovascular, blood,  
 PT electrolyte imbalance or neoplastic disorders, autoimmune diseases,  
 PT cancers  
 XX  
 PS Claim 1; SEQ ID No 33; 506pp; English.  
 XX  
 CC The invention relates to novel nucleic acids and polypeptides useful for  
 CC diagnosing, treating, preventing and/or prognosing disorders related to  
 CC these polypeptides. The polynucleotides are especially useful in the  
 CC diagnosis, prognosis, prevention and/or treatment of diseases which  
 CC include kidney disorders (e.g. renal failure or nephritis),  
 CC cardiovascular disorders (e.g. hypertension or myocardial infarction),  
 CC blood disorders (e.g. anaemia or blood coagulation disorders),  
 CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),  
 CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune  
 CC diseases, cancers, inflammatory diseases, reproductive system

CC disorders, endocrine disorders, neural activity and neurological  
 CC disorders, wound healing and respiratory disorders. AAS30165-AAS30251  
 CC represent the novel human renal and cardiovascular-associated nucleic  
 CC acid sequences of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at:  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 92.6%; Score 438.8; DB 22; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCCAGATTATTCAGTGGCAATTTCTTGCTATTAATTTGGCCGATTAATGATCACTA 60  
 DB 2 GCCCAGATTATTCAGTGGCAATTTCTTGCTATTAATTTGGCCGATTAATGATCACTA 61  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATTAACAGACTGAA 120  
 DB 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATTAACAGACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTCTTTATAGTATT 180  
 DB 122 ATACGGAATCAAGTTAAAAAGAGATGATCCTTGCCAAACGTTTTCTTTATAGTATT 181  
 QY 181 ACTGATGATTAATGCTGATGCCATTTTGTGAGGAACCTCTTACCTGCTCAGGTA 240  
 DB 182 ACTGATGATTAATGCTGATGCCATTTTGTGAGGAATTTCTTCACTGCTCAGGTA 241  
 QY 241 GAATATCAGGATACCATTAACCTCTTGCTAGTGAATGTTATTCGCAATTAAAGATGCT 300  
 DB 242 GAATATCAGGATACCATTAACCTCTTGCTAGTGAATTTTATTCGCCATTAACAGTCT 301  
 QY 301 TTGAACCCATTTCTTATCTCTGACCAAGACCATTTAAAGATGATTCATCGGTTT 360  
 DB 302 TTGAACCCATTTCTTATCTCTGACCAAGACCATTTAAAGATGATTCATCGGTTT 361  
 QY 361 TGGCATACTACAGCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 420  
 DB 362 TGGTATTACTACAGCAAGAAATCTATGAGCAGCAAGGTATCAGAAACATATGCTC 419  
 QY 421 CATCATTCATCTGGGGGAAATGTGACCCTGACAGAGATGCCACTGAGTTAA 474  
 DB 420 CATCATTCATCTGGGGTGAATGTGACCCTGACAGAGATGCCACTGAGTTAA 473

#### RESULT 14

AAS34845  
 ID AAS34845 standard; CDNA; 530 BP.

XX AC AAS34845;

DT 04-DEC-2001 (first entry)

XX CDNA encoding novel human neoplastic disease associated polypeptide #79.

XX Human; neoplastic disease associated polypeptide; cancer; gene therapy;  
 XX hyperproliferative disorder; neural disorder; immune system disorder;  
 XX muscular disorder; reproductive disorder; gastrointestinal disorder;  
 XX pulmonary disorder; cardiovascular disorder; renal disorder;  
 XX neuroprotective; cytoskeletal; anti inflammatory; vasotropic; ss.

OS Homo sapiens.

XX WO200155163-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01358.

XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.



XX Claim 4; SEQ ID No 89; 687bp; English.  
PS The present invention relates to the isolation of novel human neoplastic  
XX disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful in the diagnosis, treatment, prevention and/or  
CC prognosis of disorders involving neoplastic disease such as  
CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder  
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar  
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may  
CC also be useful for treating other disorders such as neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC and renal disorders. The polynucleotide sequences of the invention are  
CC also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences  
CC encoding for the novel human neoplastic disease associated polypeptides  
CC of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 92.6%; Score 438.8; DB 22; Length 530;  
Best Local Similarity 96.6%; Pred. No. 3.5e-109;  
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

QY 1 GCCGAGATTATTCAGTGGCAATTTTCTGTATTAATTTGGCCGATTTATCATCATATA 60  
Db 2 GCCGAGATTATTCAGTGGCAATTTTCTGTATTAATTTGGCCGATTTATCATCATATA 61

QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCATTAACAGAACTGAA 120  
Db 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAAGTGCATTAACAGAACTGAA 121

QY 121 ATAGGAATCAAGTTAAAAAGAGATGATCTTCCCAAAGCTTTTCTTATATGATTT 180  
Db 122 ATAGGAATCAAGTTAAAAAGAGATGATCTTCCCAAAGCTTTTCTTATATGATTT 181

QY 181 ACTGATGATATGCTGGATACCATTTTGTAGCAAACTCTTTCACGCTTCAGGTA 240  
Db 182 ACTGATGATATGCTGGATACCATTTTGTAGCAAACTCTTTCACGCTTCAGGTA 241

QY 241 GAAATACCAAGTACCATTAACCTTTGGGTAGTGTATTTGCCATTAAACAGTGT 300  
Db 242 GAAATACCAAGTACCATTAACCTTTGGGTAGTGTATTTGCCATTAAACAGTGT 301

QY 301 TTGAACCAATTCCTATATCTCTGACCAAAAGCATTTAAAGAAATGATTCATCGGTTT 360  
Db 302 TTGAACCAATTCCTATATCTCTGACCAAAAGCATTTAAAGAAATGATTCATCGGTTT 361

QY 361 TGGCATTAACAGCAAGAAAGAAATCTATGAGACAGCAAGAGTATCGAAACATATGCTC 420  
Db 362 TGGATTAACAGCAAGAAAGAAATCTATGAGACAGCAAGAG--TCGAAACAAATATGCTC 419

QY 421 CATATTCATCTGGGGGAAATGTGGCACTGACAGAGATGCCACATGAGTTAA 474  
Db 420 CATCATTCATCTGGGTGAAATGTGGCACTGACAGAGATGCCACATGAGTTAA 473

RESULT 15  
ABK43875  
ID ABK43875 standard; cDNA; 530 BP.  
XX  
AC ABK43875;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE DNA encoding novel central nervous system protein #455.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
hyperproliferative disorder; neoplasm; cardiovascular disorder;

KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX MO200155318-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01332.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.

[illegible]

PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251868.	
PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251889.	
PR	08-DEC-2000;	2000US-0251980.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0255976.	
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
XX	Rosen CA,	Barash SC, Ruben SW;	
XX			
DR	MP1: 2001-581633/55.		
DR	P-PSDB: AA087545.		
PT	New isolated nucleic acid encoding a protein for diagnosing,		
PT	preventing, treating or ameliorating medical conditions and used as		
XX	food additives or preservatives -		
XX			
PS	Claim 1; SEQ ID NO 465; 837bp; English.		
XX			
CC	The invention describes an isolated nucleic acid molecule (I) encoding a		
CC	novel central nervous system protein. (I) and polypeptides (II) encoded		
CC	by (I), are used to treat a medical conditions and in diagnosis of a		
CC	pathological condition. Disorders which are diagnosed or treated include		
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative		
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders		
CC	e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,		
CC	angiogenesis, nervous system disorders e.g. Alzheimer's disease and		
CC	amyotrophic lateral sclerosis, infections caused by bacteria, viruses		
CC	e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders		
CC	e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,		
CC	adenocarcinomas and irritable bowel syndrome, reproductive system		
CC	disorders e.g. testicular feminization, endocrine disorders e.g. diabetes		
CC	and pituitary dwarfism, cancers and disorders at the cellular level e.g.		
CC	leukemia, disorders involving neovascularisation e.g. malignancies,		
CC	respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.		
CC	acute kidney failure and blood related disorders e.g. myocardial		
CC	infarction. The polypeptides can also be used to aid wound healing and		
CC	interfacial cell proliferation, to prevent skin aging due to sunburn, to		
CC	maintain organs before transplantation, for supporting cell culture of		
CC	primary tissues, to regenerate tissues and in chemotaxis. The		
CC	polypeptides can also be used as a food additive or preservative to		
CC	increase or decrease storage capabilities, fat content, lipid, protein,		
CC			
Query Match	92.6%;	Score 438.8;	DB 23; Length 530;
Best Local Similarity	96.6%;	Pred. No. 3.5e-109;	
Matches	455; Conservative	1; Mismatches	13; Indels 2; Gaps 1.
OY	1	GCCGCAATTTATTCAGTGCAGCAATTTTCTGGATTAATTTGGCCGATTTATCATCAT	60
Db	2	GCCGAGATTATTCAGTGCAGCAATTTTCTGGATTAATTTGGCCGATTTATCATCAT	61
OY	61	GTTTTTCCATGAGAGCATGTTTATATGTTGTCATCAAGTGCATTAAGACATCGAA	120
Db	62	GTTTTTTCATGAGAGCATGTTTATATGTTGTCATCAAGTGCATTAAGACATCGAA	121
OY	121	ATACGAGATCAAGTTAAAAAAGATGATCTCTGGCCAAACGTTTTTCTTTATAGTATT	180
Db	122	ATACGAGATCAAGTTAAAAAAGATGATCTCTGGCCAAACGTTTTTCTTTATAGTATT	181
OY	181	ACTGATCATTAATGCTGATATCCCATTTTGTAGAGCAACCTCTTACCTGCTTCAGGTA	240
Db	182	ACTGATCATTAATGCTGATATCCCATTTTGTAGAGCAATTTCTTTCACGCTTCAGGTA	241
OY	241	GAATATCAGAGTATCAATACCTCTTGGGTAGTGAATGGTATTTCTGCAATTAACAATGCT	300
Db	242	GAATATCAGAGTATCAATACCTCTTGGGTAGTGAATTTTATCTGCCCAATTAACAATGCT	301
OY	301	TTGAAACCAATTCCTATATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT	360
Db	302	TTGAAACCAATTCCTATATCTCTGACCAAGACCATTTAAAGAAATGATTCATCGGTTT	361

QY 361 TGGATTAATACAGACAAAGAAATCTATGACAGCAAGTATCAGAAAATATGCTC 420  
 Db 362 TGGATTAATACAGACAAAGAAATCTATGACAGCAAGTATCAGAAAATATGCTC 419  
 QY 421 CATCATTCATCTGGGGGAAATGTGGCCACTGAGAGATGCCACTGAGTTAA 474  
 Db 420 CATCATTCATCTGGGGGAAATGTGGCCACTGAGAGATGCCACTGAGTTAA 473

Search completed: February 3, 2003, 13:01:45  
 Job time : 231 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2003, 12:57:12 / Search time 1940 Seconds  
(without alignments)  
3957.042 Million cell updates/sec

Title: US-10-049-568-1

Perfect score: 474

Sequence: 1 gccacgattatcagtgcc.....ggagatgccacctgagttaa 474

Scoring table:

IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

EST.\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_ochei:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	320.4	67.6	841	14	BQ228832 AGENCOURT
2	136.4	28.8	744	12	BG304121 F130F05.x
3	98.4	20.8	235	10	AM436170 75600 MAR
4	87.6	18.5	197	12	BF942735 EST-CD15N
5	69.4	14.6	286	9	AI375172 fc10b12.x
6	50.6	10.7	456	12	BF021857 uys5h08.y

7	49.6	10.5	574	10	BB632910	BB632910
8	44	9.3	452	14	HL7265	HL7265 ym37b10.r1
9	42.6	9.0	570	17	AQ697633	AQ697633 HS_5525_B
10	42.4	8.9	1101	17	CNS016TX	AL107199 Drosophila
11	42.2	8.9	896	17	CNS07DBB	AL440253 T7 end of
12	41.4	8.7	1101	17	CNS0145U	AL103740 Drosophila
13	41	8.6	1101	17	CNS0039G	AL063921 Drosophila
14	40	8.4	553	17	AQ75843	AQ75843 HS_2151_B
15	39.8	8.4	195	12	BF178875	BF178875 601807574
16	39.8	8.4	833	17	CNS042M5	AL272030 Tetradon
17	39.8	8.4	1024	17	CNS020AX	AL206466 Tetradon
18	39.4	8.3	797	17	CNS003F8	AL064634 Drosophila
19	39.4	8.3	1101	17	CNS00BEL	AL069094 Drosophila
20	39.4	8.3	1101	17	CNS0182P	AL108811 Drosophila
21	39.2	8.3	602	13	BMS68930	BMS68930 K36c08.y
22	39.2	8.3	631	17	AG036218	AG036218 Pan trogl
23	39.2	8.3	900	17	CNS03M11	AL250174 Tetradon
24	39.2	8.2	622	17	BH681813	BH681813 BOMDB70TR
25	39	8.2	686	17	BH575483	BH575483 BCGEM92TF
26	38.8	8.2	356	14	C91574	C91574 C91574 Dict
27	38.8	8.2	574	14	BQ388649	BQ388649 NISC_mq03
28	38.8	8.2	691	17	BH152955	BH152955 Gm_Umb001
29	38.6	8.1	241	9	AL117864	AL117864 p8730b11
30	38.6	8.1	587	17	CNS0464P	AL276226 Tetradon
31	38.4	8.1	368	10	AV698306	AV698306 AV698306
32	38.2	8.1	643	17	AZ432873	AZ432873 IM0218A01
33	38.2	8.1	1101	17	CNS00ZOL	AL097915 Drosophila
34	38	8.0	689	17	AQ779777	AQ779777 HS_5572_B
35	38	8.0	905	17	AZ544622	AZ544622 ENTPU41TF
36	37.8	8.0	342	12	BF986555	BF986555 CM3-CN005
37	37.8	8.0	346	9	AA128093	AA128093 z114e05.r
38	37.8	8.0	376	10	AW368885	AW368885 H12-HT019
39	37.8	8.0	839	12	BF02125	BF02125 601559730
40	37.6	7.9	1101	17	CNS0172R	AL108705 Drosophila
41	37.4	7.9	395	14	BQ739689	BQ739689 PESTOap4
42	37.4	7.9	593	17	BH007158	BH007158 ee63g04.x
43	37.4	7.9	1190	17	CNS020N7	AL206908 Tetradon
44	37.2	7.8	428	17	AL760324	AL760324 Arabidops
45	37.2	7.8	512	12	BF725643	BF725643 Dxl17f03.y

#### ALIGNMENTS

RESULT 1  
BQ228832 841 bp mRNA linear EST 02-MAY-2002  
AGENCOURT 7522585 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6051251  
5', mRNA sequence.  
BQ228832  
BQ228832.1 GI:20410232  
EST.  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE  
NIH-MGC http://mgc.nci.nih.gov/.  
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC/DCMP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM13304 row: m column: 12  
High quality sequence start: 87  
High quality sequence stop: 603.  
Location/Qualifiers

#### FEATURES

```

source
1.841
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6051251"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site_1: Nct1; Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
Technology: "

BASE COUNT      240 a      160 c      170 g      269 t      2 others
ORIGIN
Query Match      67.6%; Score 320.4; DB 14; Length 841;
Best Local Similarity 95.4%; Pred. No. 1.5e-70;
Matches 330; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GCCCAGATTATTCAGCGCAATTTCTGTGATTATTTGGCCGATTTATCATCAT 60
305 GCCCAGATTATTCAGCGCAATTTCTGTGATTATTTGGCCGATTTATCATCAT 364

QY 61 GTTTTTCCTATGAGCAGTATTTTATGATGTTTCATCAAGTCCATTAACAGCACTGAA 120
365 GTTTTTCCTATGAGCAGTATTTTATGATGTTTCATCAAGTCCATTAACAGCACTGAA 424

QY 121 ATACGGAATCAAGTTAAAGAGATGATCTTGCACAAAGTTTCTTTATGATATT 180
425 ATACGGAATCAAGTTAAAGAGATGATCTTGCACAAAGTTTCTTTATGATATT 484

QY 181 ACGATGATTAATGCTGGATACCATTTTGTGCGCAACCTGTTCACTGCTTCAGGTA 240
485 ACGATGATTAATGCTGGATACCATTTTGTGCGCAACCTGTTCACTGCTTCAGGTA 544

QY 241 GAATACCAAGTACCATTAACCTCTTGGGATGATGATTTGTTATTTCTGCCATTAACAGTCT 300
545 GAATACCAAGTACCATTAACCTCTTGGGATGATGATTTTATTTCTGCCATTAACAGTCT 604

QY 301 TTGAACCAATTCCTTAATCTGACCAACCAAGCATTTTAAAGAA 346
605 TTGAACCAATTCCTTAATCTGACCAACCAAGCATTTTAAAGAA 650

RESULT 2
BG304121/c      744 bp      mRNA      linear      EST 23-FEB-2001
LOCUS      f130f05.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
DEFINITION      315552 3' similar to TR:Q9VBP0 Q9VBP0 CG5042 PROTEIN.; mRNA
SEQUENCE
BG304121
VERSION      BG304121.1 GI:13101648
KEYWORDS      EST.
SOURCE      zebrafish.
ORGANISM      Danio rerio
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 744)
Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
Mertin,J., Pape,D., Stepien,M., Underwood,K., Theising,B., Riltter
,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU Zebrafish EST Project 1999
Unpublished (1999)
Contact: S.L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.

```

```

FEATURES
source
Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T7 from Gibco
High quality sequence stop: 510.
Location/Qualifiers
1.744
/organism="Danio rerio"
/strain="AB"
/db xref="taxon:7955"
/clone="3815552"
/issue_type="Sugano Kawakami zebrafish DRA"
/sex="mixed (one male and one female, including
unfertilized eggs)"
/dev stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="Vector: pME18S-Fl3; Site_1: DraIII (CACTGCTG);
Site_2: DraIII (CACCATCTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCTTCTG], digested and cloned into distinct DraIII
sites of the pME18S-Fl3 vector (5' site CACTGCTG, 3' site
CACCATCTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science) and
kindly donated by Dr. Koichi Kawakami. Custom primers for
sequencing: 5' end primer CTTCTGCTTAAAGCTGCG and 3' end
primer CGACTCGACGCTGACGAC."

BASE COUNT      204 a      158 c      186 g      196 t
ORIGIN
Query Match      28.8%; Score 136.4; DB 12; Length 744;
Best Local Similarity 62.1%; Pred. No. 3.9e-24;
Matches 215; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 25 TTTCTGTGATTATTTATTTGGCCGATTTATTCATCATGTTTCTTCTATGACATGTT 84
744 TTTCTGTGATTATTTATTTGGCCGATTTATTCATCATGTTTCTTCTATGACATGTT 685

QY 85 TTTATGTTTCATCAAGTGCATTAACAGCACTGAAATAGGAATCAAGTTAAAGAAAG 144
684 TTTATGTTTCATCAAGTGCATTAACAGCACTGAAATAGGAATCAAGTTAAAGAAAG 625

QY 145 ATGATCTTGGCCAAAGTTTCTTTTATGATTTTATGATTTTATGATTTATGCTGATACCC 204
624 ATGATCTTGGCCAAAGTTTCTTTTATGATTTTATGATTTTATGATTTTATGCTGATACCC 565

QY 205 ATTTTGTAGCGAAACCTCTTCTTCACTGCTTCAAGTGAAGAAATCCAGATCAATACCTT 264
564 ATTTTGTAGCGAAACCTCTCTCTTAATGAGAGGTGAGAAATCCAGATCAATACCTT 505

QY 265 TGGGTAGATGATGTTATTTGCGCATTAACAGTGGCTTGAACCCATCTCTATATCTGTG 324
504 TGGGTAGATGATGTTATTTGCGCATTAACAGTGGCTTGAACCCATCTCTATATCTGTG 445

QY 325 ACCCAAGACCATTTAAAGAAATGATTCATGCGTTTGGCATTA 370
444 ACCCAAGACCATTTAAAGAAATGATTCATGCGTTTGGCATTA 399

RESULT 3
AM436170      235 bp      mRNA      linear      EST 09-JUL-2000
LOCUS      AM436170
DEFINITION      75600 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION      AM436170
VERSION      AM436170.1 GI:6971476
KEYWORDS      EST.
SOURCE      pig.
ORGANISM      Sus scrofa
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
Fahnenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casae,E.,
Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.W.

```



```

1. .286
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cbone="IMAGE:2063423"
/clone_1ib="Scaree NBHMPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below) ; Vector: pTTT3D-Pac
(pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBH, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

```

Query Match	14.6%	Score 69.4;	DB 9;	Length 286;
Best Local Similarity	98.6%;	Pred. No.3.3e-07;		
Matches 70; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	404	TCGAAAAACATATCTCCATCATTCCTGGGGGGAATGAGCCACATGACAGAGATGCC	463
QY	277	TCGAAAAACATATCTCCATCATTCCTGGGGGGAATGAGCCACATGACAGAGATGCC	218
Db	464	ACCTGAGTTAA	474
QY	217	ACCTGAGTTAA	207

RESULT 6					
BF021857					
LOCUS	BF021857	456 bp	mRNA	linear	EST-29-DEC-2000
DEFINITION	uy581808.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone				
	IMAGE:3663807.5				
	similar to TR:Q9YVG0 Q9YVG0 CGA187 PROTEIN. ; mRNA sequence.				

```

120 a      110 c      111 g      115 t
BASE COUNT
ORIGIN
/clone_l1b="McCarrey Eddy round spermatid"
/sex="male"
/tissue_type="round spermatids, pooled from multiple mice"
/dev_stage="60 day"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene
); Site 1: XhoII; Site 2: EcoRI; cDNA oligo dt-primed
[5'-(GA)10-ACTAGCTCTGAGTTCCTTTTTCCTT-3'] and directionally
cloned using 5' linkers 5'-AATTCGGCAGG-3' and
5'-CTCTGCGCG-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from lambda-unizp-R) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 98.5% recombinants.
References: J. Andiol. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D.
(Southwest Foundation for Biomedical Research, Dept. of
Genetics); excision done by E.M. Eddy, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences). Original lambda-based library is
available through ATCC, catalog #63423."
```

Query Match	10.7%	Score 50.6	DB 12	Length 456
Best Local Similarity	59.3%	Pred. No. 0.018		
Matches 86	Conservative 0	Mismatches 59	Indels 0	Gaps 0

QY 224 TTTCAGCTGCTCAGGTGAAATTAACAGATGACATTAACCTCTGTGGGTGATGATGGTAAT 283

Db 11 TTGCACGAGTGGGGACAGGGAAGTTAAGCAATACATCTCTGGATGTGGTTTTTTTCC 70

QY 284 CTGCACATTAACAGAGCTTTGAACCAATTCCTCTACTGTGACACAAAGACATTTAAG 343

Db 71 TTCCGGTAAACAGCGCCCTTAATACCCATCCTCTACACTGACGACCTCCTTTTTTAAG 133

QY 344 AATGATTCATCGGTTTTGGCATTA 368

Db 131 ACAAGTTAAACAGTGTCTGCACAA 155

RESULT 7	LOCUS	DEFINITION
BB632910	574 bp mRNA	linear EST 26-OCT-2001
BB632910	RIKEN full-length enriched, adult male	hypothalamus Mus
musculus cDNA A230094D06 5',	mRNA sequence.	

JOURNAL	The MASHU-NCI Mouse EST Project 1999
COMMENT	Unpublished (1999) Contact: Marra M/MASHU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wuston.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. WGI:1424575 Seq primer: Primer name ambiguous High quality sequence stop: 386. Location/Qualifiers 1. 456 /organism="Mus musculus" /strain="CD-1" /db_xref="taxon:10090" /clone="IMAGE:3663807"
FEATURES	source

M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga,  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
 10 (11), 1757-1771 (2000)  
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa,  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 574  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="A230094D06"  
 /clone\_lib="RIKEN full-length enriched, adult male  
 hypothalamus"  
 /sex="male"  
 /class\_type="hypothalamus"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 20.0 and subtraction to Rot = 459.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGAGATCTCCAGATTAAATTAATTCCTCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified phagescript KS(+) after bulk excision from Lambda  
 FLC 1."

BASE COUNT 153 a 139 c 141 g 141 t  
 ORIGIN

## Query Match

Best Local Similarity 10.5%; Score 49.6; DB 10; Length 574;  
 Matches 76; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 249 AGGTACATAACCTCTGGTAGTGTATTTGTCATTAAAGTCTTTGAACCC 308  
 DB 66 AGGACATACATCTCGATCGTGTCTTTTCCGCGTGAACGCCCTTAACCC 125  
 QY 309 AATTCTTAACTCTGACCAAGACATTAAAGAAATGATTCATGCTTTGGCATTA 368  
 DB 126 CATCTCTACACTCTGACGACTCTTTTAAAGACAAAGTTGAACAGTTGTGCACAA 185

## RESULT 8

LOCUS H17265 450 bp mRNA linear EST 29-JUN-1995  
 DEFINITION YM37D10.r1 Soares infant brain INIB Homo sapiens cDNA clone  
 IMAGE:50195 5', mRNA sequence.

ACCESSION H17265  
 VERSION H17265.1 GI:883505  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 450)  
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,  
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston,  
 R., Williamson, A., Woldmann, P. and Wilson, R.  
 The MASHU-Merck EST Project  
 Unpublished (1995)  
 CONTACT: Wilson RK  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 1886  
 High quality sequence stops: 396  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
 Insert length: 1886 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 396.

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 450  
 /organism="Homo sapiens"  
 /db\_xref="GDB:423005"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:50195"  
 /clone\_lib="Soares infant brain INIB"  
 /sex="female"  
 /dev\_stage="73 days post natal"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: whole brain; Vector: lambdafmd BA; Site 1: Not  
 I; Site 2: Hind III; 1st strand cDNA was primed with a Not  
 I - oligo(dT) primer [5'  
 AACTGGAAGATTCGCGCGCAGAGAAATTTTTTTTTTTTTTTT 3'];  
 double-stranded cDNA was ligated to Hind III adaptors  
 (Pharmacia), digested with Not I and directionally cloned  
 into the Not I and Hind III sites of the lambdafmd BA vector.  
 Library went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 127 a 79 c 79 g 164 t  
 ORIGIN

## Query Match

Best Local Similarity 9.3%; Score 44; DB 14; Length 450;  
 Matches 122; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 8 TTTATTCAGTGGCAATTTTCTTGATTAATTTGGCGCATTTATCATCATGTTT 67  
 DB 105 TTTGTTTGGGGTGTTCATATTTGATTTTTCCTCGATCATCTCTAAATGTTT 164  
 QY 68 CCTATGGAAGCATGTTTATAGTGTTCATCAAGTGCATACGCACTGAATACGGA 127  
 DB 165 CGGTGACATCTTGTCTTTGTTTAAATGATCAAAAGAAACAGAAATTCACGA 224  
 QY 128 ATCAAGTTAAAAAGATGATCTTGGCAACGTTTCTTATATGATTTTACGATG 187  
 DB 225 GGGCAATCTAAAGGAAAAATCTACACTCTTTTACTACTTTGATTTATTTCTCATTT 284  
 QY 188 CATTAAGCTGATACCATTTTGTGAGCAAACTCTTTCACGCTTCAGATGAAATAC 247  
 DB 285 TTGGAAAAAGAAATCTTAATGCTACTAGAAATCTTCTTCAGTTTAAAGATATTG 344  
 QY 248 CAGGTACCATAA 259  
 DB 345 GATAAACCTCA 356

RESULT 9  
LOCUS A0697633  
DEFINITION HS\_5525\_B2\_H10\_T7A\_RPCI-11 Human Male BAC Library Homo sapiens  
ACCESSION A0697633  
VERSION A0697633.1 GI:5387881  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
COMMENT 99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@edj.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research Genetics (info@resgen.com). BAC end Web Server:  
http://www.lnsc.washington.edu  
Plate: 1101 row: P column: 20  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 572.  
Location/Qualifiers  
1..572  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate=1101 Col=20 Row=P"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"  
BASE COUNT 188 a 111 c 84 g 178 t 11 others  
ORIGIN  
Query Match 9.0%; Score 42.6; DB 17; Length 572;  
Best Local Similarity 56.9%; Pred. No. 1.8;  
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

LOCUS CNS016TX 1101 bp DNA linear GSS 26-JUL-1999  
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC  
BACN16A06 of DrosBAC library from Drosophila melanogaster (fruit  
fly) genomic survey sequence.  
ACCESSION AL107199  
VERSION AL107199.1 GI:5626132  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster  
REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
AUTHORS 1 (bases 1 to 1101)  
TITLE Genoscope.  
JOURNAL Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT Determination of this BAC-end sequence was carried out as part of a  
collaboration with the European Drosophila Genome Project (EDGP) -  
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC  
library (Dros BAC) was made by Alain Billaud at CEPH (Centre  
d'Etude du Polymorphisme Humain) with funding provided by a MRC  
project grant. The DNA was prepared from embryos by Alain Bucheton  
and Genevieve Payan. It has been constructed in the vector  
pBelBAC11.  
FEATURES  
source  
1..1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACN16A06"  
/clone\_lib="DrosBAC"  
/plasmid="pBelBAC11"  
/note="end : SP6"  
BASE COUNT 225 a 178 c 59 g 330 t 309 others  
ORIGIN  
Query Match 8.9%; Score 42.4; DB 17; Length 1101;  
Best Local Similarity 29.5%; Pred. No. 2;  
Matches 112; Conservative 81; Mismatches 184; Indels 3; Gaps 1;  
Matches 112; Conservative 81; Mismatches 184; Indels 3; Gaps 1;  
OY 8 TTATTCAGTGGCAATTTTCTTGATTAATTGGCCGCAATTATCATCATAGTTT 67  
DB 973 TDMBSVAAABMSMHTMTTBSCTWTCTTTTMTTNTKWMNTTTTTT 914  
OY 68 CCTATGAGAGATGTTTATAGTTCATCAAGTCCATACAGCAATGAAATACGA 127  
DB 913 BBAATGATATVSSKTTADMMKTSWVAALAAASVAACSMGSTAATBSSSW 854  
OY 128 ATCAAGTAAAGAGATGATCTTCCCAAGCTTTTCTTTATAGTATTAAGT 187  
DB 853 TTCACMAAMATAMMMVSTTATABACSCAMTATTTTTTTTWTATACCTTT 794  
OY 188 CATATAGCTGAT---ACCATTTTGTAGCGAAACCTTTTCACGCTTCAGTAGAA 244  
DB 793 WAWWWAGTHTTWCASYSATATTTTCAACVGAATTAATACMCAGSSCWYTPASY 734  
OY 245 TACAGGATACATACCTCTTGGTAGTGTATTTCTGCATTAACAGTCTTGA 304  
DB 733 SAABAASAAVYAAACGCSYWMCHTTTTTTTATATMCSAAGAKTTATWTTTARB 674  
OY 305 ACCCAATTTCTATATCTCTGACCAACAGACATTTAAAGAAATGATTCATGCTTTG 364  
DB 673 ASTTTTWTWMAAATTANNANNNNNNNNGNNAAGTAGNAGNGAAATNTTGTNANG 614  
OY 365 ATAACTACAGCAAGAA 384  
DB 613 ATKRAAGAAAAAANRNGA 594

RESULT 10  
CNS016TX/c  
LOCUS CNS07DBB 896 bp DNA linear GSS 08-JUL-2001

DEFINITION	17 end of clone BD0AA011A12 of library BD0AA from strain CBS 94 of
ACCESSION	AL440253
VERSION	AL440253.1 GI:12223664
KEYWORDS	GSS.
SOURCE	Candida tropicalis
ORGANISM	Candida tropicalis
REFERENCE	Bukharova; Fungi; Ascomycota; Saccharomycetales; Saccharomyces; Eukaryota; Fungi; Ascomycota; Saccharomycetales; Candida.
AUTHORS	1 (bases 1 to 896)
TITLE	Soucier, J., Algle, M., Attiguenave, F., Blandin, G., Bolojin-Fukuhara, M., Bon, E., Brothier, P., Casargola, S., de-Montigny, J., Dujon, B., Dureux, P., Lepingle, A., Lorente, B., Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Porter, S., Saurin, W., Tekra, F., Toffano-Nioche, C., Weslowski-Louvel, M., Wincker, P. and Weissenbach, J.
JOURNAL	Genomic exploration of the hemiascomycetous yeasts: 1. A set of
MEDLINE	yeast species for molecular evolution studies
UNPUBLISHED	FBS Lett. 487 (1), 3-12 (2000)
REFERENCE	2 (bases 1 to 896)
AUTHORS	Blandin, G., Ozier-Kalogeropoulos, O., Wincker, P., Attiguenave, F. and Dujon, B.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 16. Candida tropicalis
JOURNAL	FBS Lett. 487 (1), 91-94 (2000)
MEDLINE	20584726
UNPUBLISHED	11152891
REFERENCE	3 (bases 1 to 896)
AUTHORS	Genoscope.
TITLE	Direct Submision
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Séquencage, 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail : segre@genoscope.cns.fr - Web : <a href="http://www.genoscope.cns.fr">http://www.genoscope.cns.fr</a> ]
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: <i>Saccharomyces bayanus</i> var. <i>uvaurum</i> , <i>Saccharomyces exiguus</i> , <i>Saccharomyces servazzii</i> , <i>Zygosaccharomyces rouxii</i> , <i>Saccharomyces kluyveri</i> , <i>Kluyveromyces thermotolerans</i> , <i>Kluyveromyces lactis</i> var. <i>lactis</i> , <i>Kluyveromyces marxianus</i> var. <i>marxianus</i> , <i>Pichia angusta</i> , <i>Debaryomyces hansenii</i> var. <i>hansenii</i> , <i>Pichia sorbitophila</i> , <i>Candida tropicalis</i> and <i>Yarrowia lipolytica</i> . Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers
source	1..896
	/organism="Candida tropicalis"
	/strain="CBS 94"
	/db_xref="taxon:5482"
	/clone="BD0AA011A12"
	/clone_1lb="BD0AA"
	/note="end : 17"
misc_feature	<1..135
	/note="similar to <i>Saccharomyces cerevisiae</i> ORF YAL005C [ SS41 ; heat shock protein of HSP70 family, cytosolic ] similar to <i>Saccharomyces cerevisiae</i> ORF YIL074C [ SS42 ; heat shock protein of HSP70 family, cytosolic ]"
	/evidence="not_experimental"
misc_feature	<4..>135
	/note="similar to <i>Saccharomyces cerevisiae</i> ORF YER103W [ SS44 ; heat shock protein of HSP70 family, cytosolic ]"
	/evidence="not_experimental"
misc_feature	<7..252
	/note="similar to <i>Saccharomyces cerevisiae</i> ORF YBL075C [ SS43 ; heat shock protein of HSP70 family, cytosolic ]"
	/evidence="not_experimental"
BASE COUNT	292 a 165 c 139 g 250 t 50 others
ORIGIN	
Query Match	8.9%; Score 42.2; DB 17; Length 896;
Best Local Similarity	43.2%; Fred. No. 2.3;
Matches	98; Conservative 18; Mismatches 111; Indels 0; Gaps 0;

QY	202	CCGATTTTGTGACCAAAACCTTCTTACTGCTTACAGTATGAATATACCAGTATACCAATAC	261
Db	280	YCGTATATTTTATGATATATASACTACACAGCGTGTGTTTTTTTTTAAATGTTTATTTAT	339
QY	262	TCCTGGTGTGATATGTTATCTGCACTTAACAGTCCGTTGAACCAATCTCTATACT	321
Db	340	TCCTAGGTATGGTTCTTCATGCTTTATTTATCTCTTTTACMACTGTCATATYTTY	399
QY	322	CTGACCAAGAACCTTTTAAAGAAATGTTATTCGGTTTGGATTAATCTACAGCAAA	381
Db	400	YGCCTTAAAGAAACATATCTGTATAGCTGTATATTTTAAACATTAATTAATACAT	459
QY	382	AAATCTATGACACGAAAGGTATCAGAAACATATGCTTCATCTTC	428
Db	460	AAATGTATATTAATAAATAAATAAACCTTGATATATATTAAMMCCATCAATC	506
RESULT 12			
LOCUS	CNS0145U		
DEFINITION	CNS0145U	1101 bp	DNA linear GSS 26-JUL-1999
ACCESSION	CNS0145U		
VERSION	AL103740.1	GI:5615351	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Drosophila melanogaster.		
REFERENCE	Bukacinska; Melanogaster.		
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
JOURNAL	Ephydroidea; Drosophilidae; Drosophila.		
COMMENT	<p>1 (bases 1 to 1101)</p> <p>Genoscope.</p> <p>Direct Submission</p> <p>Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)</p> <p>Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <a href="http://www.edgp.ebi.ac.uk">http://www.edgp.ebi.ac.uk</a> - . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.</p>		
FEATURES	<p>Location/Qualifiers</p> <p>1..1101</p> <p>/organism="Drosophila melanogaster"</p> <p>/db_xref="taxon:7227"</p> <p>/clone="BACN11016"</p> <p>/clone_1ib="DrosBAC"</p> <p>/plasmid="pBelOBAC11"</p> <p>/note="end : T7"</p>		
BASE COUNT	336 a	123 c	4 g 358 t 280 others
ORIGIN			
Query Match	8.7%; Score 41.4; DB 17; length 1101;		
Best Local Similarity	38.1%; Pred. No. 3.6;		
Matches	96; Conservative	9; Mismatches	147; Indels 0; Gaps 0;
QY	8	TTTATTCAGGCGAATTTCTCTGATATTAATTTGGCGCATTTATCATATGATTTT	67
Db	119	TT	178
QY	68	CCATATGACATGTTTATATAGTGTATCAAAAGTCCATACAGCAACTGAATACGA	127
Db	179	ATTTTTTTTTTATATTTTTTTTTTTTTNNNNNNNNAAAAAANNNNNANNNNNNAAA	238
QY	128	ATCAAGTAAAGAGATGATCCTGGCAACGTTTTCTTATATGATTTACGAG	187
Db	239	TTATATATTAATAAATAATTAATAATTAATTAATTAATTAATTAATTAATTTT	298

Oy	188	CATATCTGTAACCAATTTTGTGGCAAACCTTCCTACTCCTCAGTGAAATAC	247
Dn	299	TTTTTTTTTTTTTTTTTWATTNNNANNANNNNNNNNNNNNNANNNAANAAAAA	358
Oy	248	CAGGTACCATAA	259
Dn	359	NANNANAAAAAA	370
<hr/>			
RESULT 13			
CSNM0396			
LOCUS DEFINITION			
Drosophila melanogaster genome survey sequence 1ET3 end of BAC # BAOR08K10 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
AL063921			
AL063921.1 GI:4941778 GSS.			
Drosophila melanogaster. Drosophila melanogaster. Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha, Epheuroidea, Drosophilidae, Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr ) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosawa and Aaron Mamooser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.			
<hr/>			
FEATURES			
Source location/Qualifiers			
1..1101 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR08K10" /clone_1ib="RPc1-98" /note="end : 1ET3"			
<hr/>			
BASE COUNT	201 a	64 c	131 g      202 t      503 others
<hr/>			
ORIGIN			
Query Match	8.6%; Score 41;	DB 17;	Length 1101;
Best Local Similarity	15.9%; Pred. No. 4.6;	Mismatches 152;	Indels 0; Gaps 0
Matches 61;	Conservative 100;	Mismatches 152;	Indels 0; Gaps 0
<hr/>			
Oy	50	TTCATCATGATGTTTTTCCSTATGGAASCATGTTTATAGTGTATGAAGSCCATA	109
Dn	539	TAAATTMTWTTWTUUTYUWAAMATRAAMCMBAAYUHNYTUUYHUUYTUTPTMNHYMU	598
Oy	110	CAGCAACTGAAAATCGCAATCAAGTAAAAAAGAAGATGCCTTGSCCAACSGTUTTCST	169
Dn	599	THAAAHHTTNWNNTTYNAMMIMTMINTMMMAAMNHNTTUADAUUUUYUCUYYHNMHNDH	658
Oy	170	TTCATGATTTATGATGATGATATGCTGATGCCATTTGTTGAGSCAACSTTTTCAC	229
Dn	659	AHAAMAATTTTGNTHAANAATYUUYUWCAMMCSTNCJHCYUUYHHYNATHHTTHNW	718
Oy	230	TGCTTCAGGTAGAATACGAGGTACAATCACSTCTGGGTAGTAGTCTGGTATGTTGSCCA	289
Dn	719	YAHUYUUYUMAYUWMCTCTATYNHNNHHNNHKAHNTTUYAARAAAMMHNNHAYAAAAA	778

Oy	290	TTCACAGGTGGTGAACAATCTGTAATCTGTAATCTGACCAACAAGCCATTAAAGAATAA	349
Db	779	AANAATTNNHTTHNTNHNHTYMHNYMYMTCMCYCMMHYNHYATYCMWTNNHMWMTTH	838
Oy	350	TTCATCGTGTTTGGCATTAACACACAAAAGAAATGTATGACACAGCAAAGGTATCAGAA	409
Db	839	WYNNHTNNHTTTHMMNMHTTWCMWWMAATTWTWATCWAKMTMWHNNHHNNHMAC	898
Oy	410	AACATAGGCATCATCATCATCT	432
Db	899	ANHTTHNCMHNMCTCHNHT	921
RESULT 14			
AQT75843                      553 bp          DNA          linear    GSS 29-JUL-1999			
AQT75843			
LOCUS			
DEFINITION			
MS_2151_B2_B04_T7C CIT Approved Human Genomic Sperm Library D Homo			
sapiens genomic clone Plate=2151 Col=8 Row=D, DNA sequence.			
ACCESSION			
AQT75843			
VERSION			
AQT75843.1 GI:5655571			
KEYWORDS			
GSS.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE			
1 (bases 1 to 553)			
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,			
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and			
Hood,L.			
TITLE			
Sequence-tagged connectors: A sequence approach to mapping and			
scanning the human genome			
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)			
JOURNAL			
MEDLINE	99380589		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 2151 row: D column: 8 Seq primer: T7 Class: BAC ends High quality sequence stop: 553.  FEATURES source location/Qualifiers 1..553 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="Plate=2151 Col=8 Row=D" /clone_lib="CIT Approved Human Genomic Sperm Library D" /sex="male" /note="Organ: sperm; Vector: pBelobaC11; BAC Clones in E-Coli DH10B"		
BASE COUNT	230 a	90 c	89 g
ORIGIN	136 t	8 others	
Query Match	8.4%	Score 40;	DB 17; Length 553;
Best Local Similarity	55.9%	Pred. No. 8.4;	
Matches	76;	Conservative 0;	Mismatches 60; Indels 0; Gaps 0;
Oy	308	CAAATCTCTACTCTGACCAACAAGCCATTAAAGAAGATGATTCATCGATTGGGCATA	367
Db	184	CAATTTATATAGAAAAAACCTAATTAACCTATATTAATAAAGGGCAAAACATCTGAGTAGA	243
Oy	368	ACTACACACAAAGAAATCTATGACACGCAAGGTATCACAAAAATATGCTTCATCATT	427
Db	244	CATTCTTCAAAAGAAAGATATATGACACCAAGATATACGAAAAAGATGCTCAACATCATT	303
Oy	428	CATCTGGGGGGAATG	443



Db 304 GATCATAGAGAAATG 319

RESULT 15  
 BFL178875/c 195 bp mRNA linear EST 31-OCT-2000  
 LOCUS 601807574F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:4038114 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BFL178875  
 VERSION BFL178875  
 KEYWORDS BFL178875.1 GI:11057017  
 SOURCE EST.  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS NIH-MGC  
 TITLE NIH-MGC  
 JOURNAL NIH-MGC  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM9316 row: d column: 19  
 High quality sequence stop: 193.  
 Location/Qualifiers

FEATURES

source  
 1..195  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_1b="NCI CGAP Mam5"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: Not1; Cloned unidirectionally. Primer: Oligo dt.  
 library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Robin Humphreys,  
 NIH"

BASE COUNT 108 a 17 c 20 g 50 t  
 ORIGIN

Every Match 8.4%; Score 39.8; DB 12; Length 195;  
 Best Local Similarity 52.0%; Pred. No. 9.8;  
 Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
 QY 23 TTTTCTGGTATTAATGGCGCATTTATCATCATGTTTCTATGGAAGCATGT 82  
 Db 187 TTTTCTGGTATTAATGGCGCATTTATCATCATGTTTCTATGGAAGCATGT 128  
 QY 83 TTTATAGTGTTCATCAAGTGCATTAACAGCACTGAATACGAATCAAGTTAAAAAG 142  
 Db 127 GCCTAAGATTAATTTTATGAATGATGAAAGAAATAGTTCTCAAGAAAAAAT 68  
 QY 143 AGATGATCCCTGGCAAGCTTTTCTTATAGTATTTAAGTATGATGATGATAT 193  
 Db 67 AGAAATCCATTTTTCACAACTATTTGTATACATTATGACTTCATTTT 17

Search completed: February 3, 2003, 14:18:56  
 Job time : 1947 secs



•  
•  
•  
•

[illegible]

Db 229 AATTGCCATAAAAGAGATGAAAAAGTATTAGTATAAAAAATTGTCATCGATA 170  
Qy 158 AACGTTTTCTTTATAGATTATGATGATCATTTAT 193  
Db 169 AAGAGTTTTTATGATATTTTCTTACTTATCTAT 134

RESULT 2  
US-08-853-659A-6/c

; Sequence 6, Application US/08853659A  
; Patent No. 5925522  
; GENERAL INFORMATION:  
; APPLICANT: Wong, K.K.; Saffer, J.D.  
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection  
; TITLE OF INVENTION: Of A  
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Paul W. Zimmerman  
; ADDRESSEE: Intellectual Property Services  
; ADDRESSEE: Battelle Memorial Institute  
; ADDRESS: P.O. Box 999  
; STREET: Washington Way  
; CITY: Richland  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 99352  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Processor (WordPerfect 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/853,659A  
; FILING DATE: Unknown  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: none  
; FILING DATE: n/a  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8967 bases  
; TYPE: nucleotide  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
; FEATURE:  
; OTHER INFORMATION: SEQ ID NO:6 corresponds to  
; OTHER INFORMATION: nucleotides 15735 through 24701 of SEQ ID NO:2  
; DB-853-659A-6  
Query Match 8.1%; Score 38.4; DB 2; Length 8967;  
Best Local Similarity 62.5%; Pred. No. 0.37; Mismatches 36; Indels 0; Gaps 0;  
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
Qy 98 AAATGCCATAACGACGCACTGAATACGAAATCAAGTAAAAAGAGATGATCTTGCCCA 157  
Db 7124 AAATGCCATAAAAGAGATGAAAAAGTATTAGTATAAAAAATTGTCATCGATA 7065  
Qy 158 AACGTTTTCTTTATAGATTATGATGATCATTTAT 193  
Db 7064 AAGAGTTTTTATGATATTTTCTTACTTATCTAT 7029

RESULT 3  
US-08-853-659A-9

; Sequence 9, Application US/08853659A  
; Patent No. 5925522  
; GENERAL INFORMATION:  
; APPLICANT: Wong, K.K.; Saffer, J.D.  
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection  
; TITLE OF INVENTION: Of A  
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella  
; NUMBER OF SEQUENCES: 67

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: Paul W. Zimmerman  
; ADDRESSEE: Intellectual Property Services  
; ADDRESSEE: Battelle Memorial Institute  
; ADDRESS: P.O. Box 999  
; STREET: Washington Way  
; CITY: Richland  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 99352  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Processor (WordPerfect 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/853,659A  
; FILING DATE: Unknown  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: none  
; FILING DATE: n/a  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8967 bases  
; TYPE: nucleotide  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
; US-08-853-659A-9

## Query Match 8.1%; Score 38.4; DB 2; Length 8967;

Best Local Similarity 62.5%; Pred. No. 0.37; Mismatches 36; Indels 0; Gaps 0;  
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 98 AAATGCCATAACGACGCACTGAATACGAAATCAAGTAAAAAGAGATGATCTTGCCCA 157  
Db 1844 AAATGCCATAAAAGAGATGAAAAAGTATTAGTATAAAAAATTGTCATCGATA 1903  
Qy 158 AACGTTTTCTTTATAGATTATGATGATCATTTAT 193  
Db 1904 AAGAGTTTTTATGATATTTTCTTACTTATCTAT 1939

RESULT 4  
US-08-853-659A-64/c

; Sequence 64, Application US/08853659A  
; Patent No. 5925522  
; GENERAL INFORMATION:  
; APPLICANT: Wong, K.K.; Saffer, J.D.  
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection  
; TITLE OF INVENTION: Of A  
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Paul W. Zimmerman  
; ADDRESSEE: Intellectual Property Services  
; ADDRESSEE: Battelle Memorial Institute  
; ADDRESS: P.O. Box 999  
; STREET: Washington Way  
; CITY: Richland  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 99352  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Processor (WordPerfect 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/853,659A  
; FILING DATE: Unknown  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 8967 bases
;   TYPE: nucleotide
;   STRANDEDNESS: single stranded
;   TOPOLOGY: linear
US-08-853-659A-64
```

```
Query Match      8.1%; Score 38.4; DB 2; Length 8967;
Best Local Similarity 62.5%; Pred. No. 0.37; 36; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```

```
Qy 98 AAGTGCCTAATACGCACTGAATATACGAATCAAGTTAAAGAGATGATCCTTGCCA 157
Db 7124 AATATGCCATTAAGAGATGAAAAAGATATTTAGTATTAAGATCTTTCATCGATA 7065
Qy 158 AACGTTTTCTTTATAGTATTTACTGATGCATTAT 193
7064 AAGATTTTATGATATTTTCTTACTCTTATACTAT 7029
```

```
RESULT 5
US-08-853-659A-67
; Sequence 67, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 8967 bases
;   TYPE: nucleotide
;   STRANDEDNESS: single stranded
;   TOPOLOGY: linear
US-08-853-659A-67
```

```
Query Match      8.1%; Score 38.4; DB 2; Length 8967;
Best Local Similarity 38.5%; Pred. No. 0.37; 36; Indels 0; Gaps 0;
Matches 37; Conservative 23; Mismatches 36; Indels 0; Gaps 0;
```

Qy 98 AAGTGCCTAATACGCACTGAATATACGAATCAAGTTAAAGAGATGATCCTTGCCA 157  
Db 1844 AATATGCCATTAAGAGATGAAAAAGATATTTAGTATTAAGATCTTTCATCGATA 1903  
Qy 158 AACGTTTTCTTTATAGTATTTACTGATGCATTAT 193  
7064 AAGATTTTATGATATTTTCTTACTCTTATACTAT 7029

```
Db 1904 AAGAGUUUUUUAUGAUAUUUUUAUCUUAUACUUA 1939
```

```
RESULT 6
US-08-853-659A-2/C
; Sequence 2, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
```

```
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 24701 bases
;   TYPE: nucleotide
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
US-08-853-659A-2
```

```
Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.48; 36; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```

```
Qy 98 AAGTGCCTAATACGCACTGAATATACGAATCAAGTTAAAGAGATGATCCTTGCCA 157
Db 22858 AATATGCCATTAAGAGATGAAAAAGATATTTAGTATTAAGATCTTTCATCGATA 22799
Qy 158 AACGTTTTCTTTATAGTATTTACTGATGCATTAT 193
Db 22798 AAGATTTTATGATATTTTCTTACTCTTATACTAT 22763
```

```
RESULT 7
US-08-853-659A-3
; Sequence 3, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
```

```
STATE: Washington
COUNTRY: U.S.A.
ZIP: 99352
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
COMPUTER: IBM PC/XT/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (WordPerfect 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,659A
FILING DATE: Unknown
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: none
FILING DATE: n/a
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24701 bases
TYPE: nucleotide
STRANDEDNESS: double stranded
TOPOLOGY: linear
US-08-853-659A-3
```

```
Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.48;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
```

```
QY 98 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1844 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 1903
QY 158 AACGTTTCTTATAGTATTACTGATGATCAT 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1904 AAGAGTTTTTATGATATTTTACTCTTATACAT 1939
```

```
RESULT 8
US-08-853-659A-60/c
; Sequence 60, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saifer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: OF A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
```

```
TOPOLOGY: linear
US-08-853-659A-60
Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 62.5%; Pred. No. 0.48;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 98 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22858 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 22799
QY 158 AACGTTTCTTATAGTATTACTGATGATCAT 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22798 AAGAGTTTTTATGATATTTTACTCTTATACAT 22763
```

```
RESULT 9
US-08-853-659A-61
; Sequence 61, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saifer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: OF A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24701 bases
; TYPE: nucleotide
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
US-08-853-659A-61
```

```
Query Match      8.1%; Score 38.4; DB 2; Length 24701;
Best Local Similarity 38.5%; Pred. No. 0.48;
Matches 37; Conservative 23; Mismatches 36; Indels 0; Gaps 0;
QY 98 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1844 AAATGCCATAACAGCACTGAATACGAAATCAAGTAAAGATATGATCCTTGCCA 1903
QY 158 AACGTTTCTTATAGTATTACTGATGATCAT 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1904 AAGAGUUUUUUGAUAUUUUUUUUAUUUUUUUAUUUUUUUAUUUUUUUAUUUUUUUA 1939
```

```
RESULT 10
US-08-991-789A-169/c
; Sequence 169, Application US/08991789A
; Patent No. 6225054
```

GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: TREATMENT AND METHODS FOR THE  
NUMBER OF SEQUENCES: 292  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed IP Law Group  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,789A  
FILING DATE: 11-Dec-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 169:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1265 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 169:  
US-08-991-789A-169

Query Match 8.0%; Score 37.8; DB 4; Length 1265;  
Best Local Similarity 50.8%; Pred. No. 0.34; Mismatches 0; Gaps 0;  
Matches 90; Conservative 0; Indels 87; Indels 0; Gaps 0;

QY 240 AGAATACCGAGTACCAATCTTGGGATGATTGTTATCTGACATTAACAGTGC 299  
DB 792 AATATAAAGAGGCTTCAGATTATCTGCTTCATATATATTTTCTTTAAAGAAA 733

QY 300 TTGAACCAATCTCTATCTCTGACCAAGACATTTAAAGAAATGATTGCTT 359  
DB 732 ATATCAACCAATGTCATGCACTGTTTTCAAAGCATTTAAATAGAGGGTAAACCTT 673

QY 360 TTGGCATTAACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATAT 416  
DB 672 TGGAAATTATACAGAAAGAAATGATTACTTTATGCAATAAAAAATTAATATATAT 616

RESULT 11  
US-09-062-451-169/c  
Sequence 169, Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA

ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-APR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 169:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1265 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-169

Query Match 8.0%; Score 37.8; DB 4; Length 1265;  
Best Local Similarity 50.8%; Pred. No. 0.34; Mismatches 0; Gaps 0;  
Matches 90; Conservative 0; Indels 87; Indels 0; Gaps 0;

QY 240 AGAATACCGAGTACCAATCTTGGGATGATTGTTATCTGACATTAACAGTGC 299  
DB 792 AATATAAAGAGGCTTCAGATTATCTGCTTCATATATATTTTCTTTAAAGAAA 733

QY 300 TTGAACCAATCTCTATCTCTGACCAAGACATTTAAAGAAATGATTGCTT 359  
DB 732 ATATCAACCAATGTCATGCACTGTTTTCAAAGCATTTAAATAGAGGGTAAACCTT 673

QY 360 TTGGCATTAACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATAT 416  
DB 672 TGGAAATTATACAGAAAGAAATGATTACTTTATGCAATAAAAAATTAATATATAT 616

RESULT 12  
US-09-598-326-169/c  
Sequence 169, Application US/09598326  
Patent No. 6423496  
GENERAL INFORMATION:  
APPLICANT: Fridakis, Tony N.  
Smith, John M.  
Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
NUMBER OF SEQUENCES: 247  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed Intellectual Property Law Group PLLC  
STREET: 701 Fifth Avenue, Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/598,326  
FILING DATE: 20-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419D1

```
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 169:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1265 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-09-598-326-169

Query Match      8.0%; Score 37.8; DB 4; Length 1265;
Best Local Similarity 50.8%; Pred. No. 0.34;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAAATACAGGATGACATACCTCTGGTAGATGATGTTATTCGCCATTAACAGTGC 299
    |||
DB 792 AAATTAACAGGCTTCAGATATCTGGCTTCATATATATTTCTTTAAAGAAA 733
    |||
    300 TTGAACCAATTCCTATATCTGACCAAGACCAATTTAAAGAAATGATTCATCGGT 359
    |||
DB 732 ATATCAACCATTTGCAATGCACTGTTTTCAAGCATTAAATAGAGGTAAACCCCT 673
    |||
QY 360 TTGGCATTAACACAGCAAGAAATCTATGACAGCAAGGATTCAGAAACATAT 416
    |||
DB 672 TGGAAATTAATACAGAAAGAAATGATTCATTATGCAATAAATTAATTAATAT 616
    |||

RESULT 13
US-09-134-001C-1232/c
/ Sequence 1232, Application US/09134001C
/ Patent No. 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14
/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NO 1232
/ LENGTH: 1557
/ TYPE: DNA
/ ORGANISM: Staphylococcus epidermidis
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (16)
/ OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1232

Query Match      8.0%; Score 37.8; DB 4; Length 1557;
Best Local Similarity 54.7%; Pred. No. 0.35;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 337 TTTAAGAAATGATTCATCGGTTTGGCATTAACACAGCAAGAAATCTATGACAGC 396
    |||
DB 1293 TTCAACAGTAATTTACTCTTTGTGGATACACACAGCAACAGCTGATGAAACAG 1234
    |||
QY 397 AAAGTATCAGAAAAACATATGCTTCATTCATCTGGGGGAAATGTGCCACTGCAGG 456
    |||
DB 1233 AATATGATTAATCTCAATGTTTCATATTAATTTGTTCCAATGCTGAGCTCAAT 1174
    |||
QY 457 AGATGCCACCTGAGTTA 473
    |||
DB 1173 ATATGCTGATGAATCA 1157
    |||

RESULT 14
```

```
US-08-991-789A-290/c
/ Sequence 290, Application US/08991789A
/ Patent No. 6225054
/ GENERAL INFORMATION:
/ APPLICANT: Fridakis, Tony N.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ TREATMENT AND DIAGNOSIS OF BREAST CANCER
/ NUMBER OF SEQUENCES: 292
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Seed IP Law Group
/ STREET: 701 Fifth Avenue, Suite 6300
/ City: Seattle
/ STATE: Washington
/ COUNTRY: USA
/ ZIP: 98104-7092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/991,789A
/ FILING DATE: 11-Dec-1997
/ CLASSIFICATION: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Poter, Jane E. R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 210121.419C3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 622-4900
/ TELEFAX: (206) 682-6031
/ INFORMATION FOR SEQ ID NO: 290:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1646 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 290:
US-08-991-789A-290

Query Match      8.0%; Score 37.8; DB 4; Length 1646;
Best Local Similarity 50.8%; Pred. No. 0.36;
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAATACAGGATGACATACCTCTGGTAGATGATGTTATTCGCCATTAACAGTGC 299
    |||
DB 1165 AAATTAACAGGCTTCAGATATCTGGCTTCATATATATTTCTTTAAAGAAA 1106
    |||
QY 300 TTGAACCAATTCCTATATCTGACCAAGACCAATTTAAAGAAATGATTCATCGGT 359
    |||
DB 1105 ATATCAACCATTTGCAATGCACTGTTTTCAAGCATTAAATAGAGGTAAACCCCT 1046
    |||
QY 360 TTGGCATTAACACAGCAAGAAATCTATGACAGCAAGGATTCAGAAACATAT 416
    |||
DB 1045 TCGAATTAATACAGAAAGAAATGATTCATTATGCAATAAATTAATTAATAT 989
    |||

RESULT 15
US-09-062-451-290/c
/ Sequence 290, Application US/09062451
/ Patent No. 6344550
/ GENERAL INFORMATION:
/ APPLICANT: Fridakis, Tony N.
/ APPLICANT: Smith, John M.
/ APPLICANT: Reed, Steven G.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ TREATMENT AND DIAGNOSIS OF BREAST CANCER
/ NUMBER OF SEQUENCES: 297
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Seed and Berry LLP
/ STREET: 6300 Columbia Center, 701 Fifth Avenue
```



CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-Apr-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 290:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-290

Query Match 8.0%; Score 37.8; DB 4; Length 1646;

Best Local Similarity 50.8%; Pred. No. 0.36;  
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAATACCGAGTACCATTAACCTTGGGTAGTGGATTGTTATTCGCATTAACAGTGC 299  
DB 1165 AATTAACAGGCTTCAGATTATCTTGCTTCATTAATTATTTCTTTAAAGAAA 1106  
QY 300 TTGAACCCATTCCTAATCTGACCAAGACCATTTAAGAAATGATTCATCGGTT 359  
DB 1105 ATATCAACCCATTCATGATGACTGTTTCAAGCATTTAATAGAGGTAAACCTT 1046  
QY 360 TTGGATATCAAGACAGAAAGAAATCTATGACAGCAAGGATACAGAAACATAT 416  
DB 1045 TGAATATTATACAGAAAGAAATGATTCATTTATGCAATAAATAATATATATAT 989

Search completed: February 3, 2003, 13:02:57  
Job time : 62 secs



7  
.  
4

QY	1	GCCACATTATTCAGTGGCAATTTCTTGATTAATTTGGCGCAATTATCATCATATA	60
Db	792	GCCACATTATTCAGTGGCAATTTCTTGATTAATTTGGCGCAATTATCATCATATA	851
QY	61	GTTTTTCTATGAGAGCATGTTTATAGTTCATCAAGGCCATAACGAACTGAA	120
Db	852	GTTTTTCTATGAGAGCATGTTTATAGTTCATCAAGGCCATAACGAACTGAA	911
QY	121	ATAAGCAATCAAGTAAAAAAGAGATATCTTGGCAAACTTTTTCTTATAGTATTT	180
Db	912	ATAAGCAATCAAGTAAAAAAGAGATATCTTGGCAAACTTTTTCTTATAGTATTT	971
QY	181	ACTGATCATATGCTGGATATCCATTTTTTATAGCAAACTCTTTACCTGCTCAGTA	240

Db 972 ACTGATGATTTATGCTGATATACCATTTTGTATGATAATTTCTTCACTGCTTCAAGTA 1031  
QY 241 GAAATCCAGGTACATTAACCTCTTGAGTAGTATGTTATCTGCATTAACAGTGTCT 300  
Db 1032 GAAATCCAGGTACATTAACCTCTTGAGTAGTATGTTATCTGCATTAACAGTGTCT 1091  
QY 301 TTGACCCATTTCTTATATCTGACCAAGACCATTTAAGAAATGATTCGCTT 360  
Db 1092 TTGACCCATTTCTTATATCTGACCAAGACCATTTAAGAAATGATTCGCTT 1151  
QY 361 TGGCATTTACAGACAAAGAAATCTATGACAGCAAGGTATCGAAACATATGCTC 420  
Db 1152 TGGTATTTACAGACAAAGAAATCTATGACAGCAAGGTATCGAAACATATGCTC 1209  
QY 421 CATCATTTCTGAGGAGGAAATGTGCGCATGACAGAGATGCCACTGAGTTAA 474  
Db 1210 CATCATTTCTGAGGAGGAAATGTGCGCATGACAGAGATGCCACTGAGTTAA 1263

ULT 2  
09-989-442-33  
Sequence 33, Application US/09989442  
Publication No. US20030013649A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PJ208  
CURRENT APPLICATION NUMBER: US/09/989,442  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/225,757  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/226,868  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/216,647  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,267  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/216,880  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,270  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/251,869  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/235,834  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/234,274  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/234,223  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/228,924  
PRIOR FILING DATE: 2000-08-30  
PRIOR APPLICATION NUMBER: 60/224,518  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,369  
PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/224,519  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,964  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/241,809  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/249,299  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/236,327  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/241,785  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/244,617  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/225,268  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,368  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/251,856  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/251,868  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/229,344  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/234,997  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: 60/229,343  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,345  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,287  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,513  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/231,413  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/229,509  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/236,367  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/237,039  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,038  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/236,370  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/236,802  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,037  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,040  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/240,960  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/239,935  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/239,937  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/241,787  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246,474  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/246,532  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/249,216  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,210  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/226,681  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225,759  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/225,213

PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/227,182  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,214  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/235,836  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: 60/230,438  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/215,135  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: 60/225,266  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/249,218  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,208  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,213  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,212  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,207  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,245  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,244  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,217  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,211  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,215  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,264  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,214  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,297  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/232,400  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/231,242  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,081  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,080  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,414  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,244  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/233,064  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/233,063  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,397  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,399  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,401  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/241,808  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,826  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,786  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,221  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,475  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/231,243  
 PRIOR FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: 60/233,065  
 PRIOR FILING DATE: 2000-09-14  
 Query Match 92.6%; Score 438.8; DB 9; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 2.4e-108;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCGAGTTTATTCAGTGGCAATTTCTTGGTATTAATTTGGCCGATTTATCATGATA 60  
 DB 2 GCCGAGTTTATTCAGTGGCAATTTCTTGGTATTAATTTGGCCGATTTATCATGATA 61  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGAGTCCATAACGCACTGAA 120  
 DB 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGAGTCCATAACGCACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAGAAAGATGATCCTTGCCAAACGTTTTCTTATAGTATTT 180  
 DB 122 ATACGGAATCAAGTTAAAGAAAGATGATCCTTGCCAAACGTTTTCTTATAGTATTT 181  
 QY 181 ACTGATGATTAAGTGGATACCAATTTTGTAGGAAACCTCTTCACCTGTCAGGTA 240  
 DB 182 ACTGATGATTAAGTGGATACCAATTTTGTAGGAAATTCCTTCACCTGTCAGGTA 241  
 QY 241 GAAATACAGTACCATTAACCTCTTGGTATGATTTGTTATTCGCAATTAACAGTCT 300  
 DB 242 GAAATACAGTACCATTAACCTCTTGGTATGATTTTATTCGCAATTAACAGTCT 301  
 QY 301 TTGAACCAATTTCTTAATCTGTACCAACAGACATTTAAAGAAATGATTCATGCTTT 360  
 DB 302 TTGAACCAATTTCTTAATCTGTACCAACAGACATTTAAAGAAATGATTCATGCTTT 361  
 QY 361 TGGCAATTAACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATCATTTGCTC 420  
 DB 362 TGGCAATTAACAGCAAGAAATCTATGACAGCAAGATGATCAAGAAATCATTTGCTC 419  
 QY 421 CATCATTCATCTGGGGGAGAAATGTGGCCACTGACAGAGATCCCACTGATTTAA 474  
 DB 420 CATCATTCATCTGGGGGAGAAATGTGGCCACTGACAGAGATCCCACTGATTTAA 473  
 RESULT 3  
 US-09-764-853-137  
 ; Sequence 137, Application US/09764853  
 ; Patent No. US20020090672A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P1206  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PAM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 939  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 137  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-853-137  
 Query Match 92.6%; Score 438.8; DB 10; Length 530;  
 Best Local Similarity 96.6%; Pred. No. 2.4e-108;  
 Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 GCCGAGTTTATTCAGTGGCAATTTCTTGGTATTAATTTGGCCGATTTATCATGATA 60  
 DB 2 GCCGAGTTTATTCAGTGGCAATTTCTTGGTATTAATTTGGCCGATTTATCATGATA 61  
 QY 61 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGAGTCCATAACGCACTGAA 120  
 DB 62 GTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAGAGTCCATAACGCACTGAA 121  
 QY 121 ATACGGAATCAAGTTAAAGAAAGATGATCCTTGCCAAACGTTTTCTTATAGTATTT 180  
 DB 122 ATACGGAATCAAGTTAAAGAAAGATGATCCTTGCCAAACGTTTTCTTATAGTATTT 180

```
Db 122 ATACGAATCAAGTAAAAAGAGATGATCCTTGCCAAAGCTTTTCTTATAGATTT 181
Qy 181 ACTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAACCTTTTCACTGCTTCAGGTA 240
Db 182 ACTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAATTTCTTTCACGCTTCAGGTA 241
Qy 241 GAATATACAGGTACCATTAACCTTTGGGTAGTGAATGGTATTCGTGCATTAACAGGCT 300
Db 242 GAATATACAGGTACCATTAACCTTTGGGTAGTGAATTTTATTTCTGCGCATTAACAGGCT 301
Qy 301 TTGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 360
Db 302 TTGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 361
Qy 361 TGGCATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 420
Db 362 TGGTATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 419
Qy 421 CATCATTCATCTGGGGGAAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 474
Db 420 CATCATTCATCTGGGGTGAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 473
```

## RESULT 4

```
US-09-764-877-986
; Sequence 986, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 986
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-986
```

Query Match 92.6%; Score 438.8; DB 10; Length 530;

Best Local Similarity 96.6%; Pred. No. 2.4e-108; Indels 2; Gaps 1;

Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

```
Qy 1 GCCCAGATTTTATTCAGTGGCAATTTTCTTGATTTATTTGGCCGCAATTTATCATATA 60
Db 2 GCCCAGATTTTATTCAGTGGCAATTTTCTTGATTTATTTGGCCGCAATTTATCATATA 61
Qy 61 GTTTTTCCTATGAGAGCATGTTTATNAGTGTTCATCAAAGTGCATTAACGCACTGAA 120
Db 62 GTTTTTCCTATGAGAGCATGTTTATNAGTGTTCATCAAAGTGCATTAACGCACTGAA 121
Qy 121 ATACGATCAAGTAAAGAAAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 180
Db 122 ATACGATCAAGTAAAGAAAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 181
Qy 181 ACTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAACCTTTTCACTGCTTCAGGTA 240
Db 182 ACTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAATTTCTTTCACGCTTCAGGTA 241
Qy 241 GAATATACAGGTACCATTAACCTTTGGGTAGTGAATGGTATTCGTGCATTAACAGGCT 300
Db 242 GAATATACAGGTACCATTAACCTTTGGGTAGTGAATTTTATTTCTGCGCATTAACAGGCT 301
Qy 301 TTGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 360
Db 302 TTGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 361
Qy 361 TGGCATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 420
Db 362 TGGTATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 419
```

```
Qy 421 CATCATTCATCTGGGGGAAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 474
Db 420 CATCATTCATCTGGGGTGAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 473
```

## RESULT 5

```
US-09-895-686-37/c
; Sequence 37, Application US/09895686
; Patent No. US2002010655A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN GPCR PROTEINS
; FILE REFERENCE: PC-0044 CIP
; CURRENT APPLICATION NUMBER: US/09/895,686
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Inocyte ID No. US2002010655A1 2488822X308B1
; NAME/KEY: unsure
; LOCATION: 561
; OTHER INFORMATION: a, t, c, g, or other
US-09-895-686-37
```

Query Match 92.1%; Score 436.6; DB 10; Length 612;

Best Local Similarity 96.6%; Pred. No. 9.6e-108; Indels 2; Gaps 1;

Matches 457; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

```
Qy 2 CCCAGATTTTATTCAGTGGCAATTTTCTTGATTTATTTGGCCGCAATTTATCATATA 61
Db 497 CCCAGATTTTATTCAGTGGCAATTTTCTTGATTTATTTGGCCGCAATTTATCATATA 61
Qy 62 TTTTTCCTATGAGAGCATGTTTATNAGTGTTCATCAAAGTGCATTAACGCACTGAA 121
Db 437 TTTTTCCTATGAGAGCATGTTTATNAGTGTTCATCAAAGTGCATTAACGCACTGAA 121
Qy 122 TACGATCAAGTAAAGAAAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 181
Db 377 TACGATCAAGTAAAGAAAGATGATCCTTGCCAAAGCTTTTCTTATAGTATTT 181
Qy 182 CTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAACCTTTTCACTGCTTCAGGTA 241
Db 317 CTGATGATTAATGCTGGAATACCCATTTTGTAGCGAAATTTCTTTCACGCTTCAGGTA 241
Qy 242 AATATACAGTACCAATTAACCTTTGGGTAGTGAATGGTATTCGTGCATTAACAGGCT 301
Db 257 AATATACAGTACCAATTAACCTTTGGGTAGTGAATGGTATTCGTGCATTAACAGGCT 301
Qy 302 TGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 361
Db 197 TGAACCAATTTCTTACTCTGACCAACAAGCCATTTTAAAGAAATGATTCATCGGTTT 361
Qy 362 GGCATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 421
Db 137 GGCATTAATCAGACAAAGAAATCTATGACAGCAAGGATTCAGAAACATATGCTC 421
Qy 422 ATCATTCATCTGGGGGAAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 474
Db 79 ATCATTCATCTGGGGTGAATGTGGCACTGACAGAGATGCCACCTGAGTTAA 474
```

## RESULT 6

```
US-09-928-175-11
; Sequence 11, Application US/09928175
```

```
/ Patent No. US20020123618A1
/ GENERAL INFORMATION:
/ APPLICANT: Paszky, Christopher J.
/ APPLICANT: Gong, Jianhua
/ APPLICANT: Daugherty, Betsy
/ APPLICANT: Rogers, No. US20020123618A1ma
/ TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
/ FILE REFERENCE: 00-1229
/ CURRENT APPLICATION NUMBER: US/09/928,175
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/224,455
/ NUMBER OF SEQ ID NOS: 42
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 11
/ LENGTH: 2049
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(2046)
/ NAME/KEY: s19_peptide
/ LOCATION: (1)..(108)
US-09-928-175-11
```

Query Match 36.2%; Score 171.8; DB 10; Length 2049;

Best Local Similarity 67.4%; Pred. No. 1.1e-36; Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```
QY 10 TATTCAGTGGCAATTTTCTTGATTAATTTGGCCGATTTATCATCATGATTTTTC 69
DB 1549 TATTCCTTGAATTTTCTAGAGTGAATCTGCTGCTTTCTCATCTGTTTCC 1608
QY 70 TATGAGAGATGTTTATAGTGTTCATCAAGGCGATACAGCACTGAATAGGAT 129
DB 1609 TATATTCTATGTTCTGTTCCATTAACCAACCGCTTCAGACACAGAAATGAAT 1668
QY 130 CAAGTAAAAAGAGATGATCTTGCACAAAGTTTCTTATAGTATTTACTGATCA 189
DB 1669 TGTTTGAGAGAGAGTGGCTGTGTGCAAAATCGTTCTTTTATAGTGTCTGATGCC 1728
QY 190 TTAATGCTGATACCATTTTGTAGGAAAACCTTTTCACTGCTTCAGGTAGAAATACCA 249
DB 1729 ATCTGCTGAGATCTCTGATTTAGTAAATCTTTCCCTTCGCGGTGGAATATACCA 1788
QY 250 GGTACCATTAACCTTGGGAGTGAATGTTGTTATCTGCATTAACAGTCTTTGAACCA 309
DB 1789 GACACATGACTTCCGATGATGATTTTTCCTTCAGTTAAGAGTCTTTGAATCA 1848
QY 310 ATTCTCTATACCTTGACACACAGACATTTAAAGAAATGATTCGTTTGGCATTA 368
DB 1849 ATCTCTATACCTCTCAACACCACTTTTAAAGACAGTGAACAGCTGCTGACAA 1907
```

## RESULT 7

```
US-09-928-175-6
/ Sequence 6, Application US/09928175
/ Patent No. US20020123618A1
/ GENERAL INFORMATION:
/ APPLICANT: Paszky, Christopher J.
/ APPLICANT: Gong, Jianhua
/ APPLICANT: Daugherty, Betsy
/ APPLICANT: Rogers, No. US20020123618A1ma
/ TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
/ FILE REFERENCE: 00-1229
/ CURRENT APPLICATION NUMBER: US/09/928,175
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/224,455
/ NUMBER OF SEQ ID NOS: 42
/ SOFTWARE: Patentin Ver. 2.0
```

```
/ SEQ ID NO 6
/ LENGTH: 2193
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(2190)
/ NAME/KEY: s19_peptide
/ LOCATION: (1)..(108)
US-09-928-175-6
```

Query Match 36.2%; Score 171.8; DB 10; Length 2193;

Best Local Similarity 67.4%; Pred. No. 1.1e-36; Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

```
QY 10 TATTCAGTGGCAATTTTCTTGATTAATTTGGCCGATTTATCATCATGATTTTTC 69
DB 1693 TATTCCTTGAATTTTCTAGAGTGAATCTGCTGCTTTCTCATCATGTTTTC 1752
QY 70 TATGAGAGATGTTTATAGTGTTCATCAAGGCGATACAGCACTGAATAGGAT 129
DB 1753 TATATTCTATGTTCTGTTCCATTAACCAACCGCTTCAGACACAGAAATGAAT 1812
QY 130 CAAGTAAAAAGAGATGATCTTGCACAAAGTTTCTTATAGTATTTACTGATCA 189
DB 1813 TGTTTGAGAGAGTGGCTGTGTGCAAAATCGTTCTTTTATAGTGTCTGATGCC 1872
QY 190 TTAATGCTGATACCATTTTGTAGGAAAACCTTTTCACTGCTTCAGGTAGAAATACCA 249
DB 1873 ATCTGCTGAGATCTCTGATTTAGTAAATCTTTCCCTTCGCGGTGGAATATCA 1932
QY 250 GGTACCATTAACCTTGGGAGTGAATGTTTATCTGCATTAACAGTCTTTGAACCA 309
DB 1933 GACACATGACTTCCGATGATGATTTTTCCTTCAGTTAAGAGTCTTTGAATCA 1992
QY 310 ATTCTCTATACCTTGACACACAGACATTTAAAGAAATGATTCGTTTGGCATTA 368
DB 1993 ATCTCTATACCTCTCAACCACTTTTAAAGACAGTGAACAGCTGCTGACAA 2051
```

## RESULT 8

```
US-09-928-175-1
/ Sequence 1, Application US/09928175
/ Patent No. US20020123618A1
/ GENERAL INFORMATION:
/ APPLICANT: Paszky, Christopher J.
/ APPLICANT: Gong, Jianhua
/ APPLICANT: Daugherty, Betsy
/ APPLICANT: Rogers, No. US20020123618A1ma
/ TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
/ FILE REFERENCE: 00-1229
/ CURRENT APPLICATION NUMBER: US/09/928,175
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/224,455
/ NUMBER OF SEQ ID NOS: 42
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 2265
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(2262)
/ NAME/KEY: s19_peptide
/ LOCATION: (1)..(108)
US-09-928-175-1
```

Query Match 36.2%; Score 171.8; DB 10; Length 2265;

Best Local Similarity 67.4%; Pred. No. 1.1e-36; Matches 242; Conservative 0; Mismatches 117; Indels 0; Gaps 0;





NUMBER OF SEQ ID NOS: 340  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 169  
 LENGTH: 1265  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-924-400-169

Query Match  
 Best Local Similarity 50.8%; Score 37.8; DB 9; Length 1265;  
 Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

240 AGAATACGAGTACCATTAACCTTGGGTAGTGTGATTTCTGCAATTAACAGTGC 299  
 Db 792 AAATTAACAGGCTTCAGATTACTTGGCTTCATTAATTAATTTTCTTTAAAGAAA 733  
 Oy 300 TTGAACCCATTTCTTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGTT 359  
 Db 732 ATATCAACCATTTGTCAATGACCTGTTTTCAAAGCATTTAAAGAGGTAAACCCCTT 673  
 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATAT 416  
 Db 672 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATTAATAATATATAT 616

RESULT 12  
 US-09-810-936-169/c  
 Sequence 169, Application US/09810936  
 Patent No. US20020068285A1  
 GENERAL INFORMATION:  
 APPLICANT: Fridakis, Tony N.  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Smith, John M.  
 APPLICANT: Mishner, Linda E.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Day, Craig H.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 TITRE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER  
 FILE REFERENCE: 210121.419C11  
 CURRENT APPLICATION NUMBER: US/09/810.936  
 NUMBER OF SEQ ID NOS: 334  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 169  
 LENGTH: 1265  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-810-936-169

Query Match  
 Best Local Similarity 50.8%; Score 37.8; DB 10; Length 1265;  
 Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

240 AGAATACGAGTACCATTAACCTTGGGTAGTGTGATTTCTGCAATTAACAGTGC 299  
 Db 792 AAATTAACAGGCTTCAGATTACTTGGCTTCATTAATTAATTTTCTTTAAAGAAA 733  
 Oy 300 TTGAACCCATTTCTTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGTT 359  
 Db 732 ATATCAACCATTTGTCAATGACCTGTTTTCAAAGCATTTAAAGAGGTAAACCCCTT 673  
 Oy 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATAT 416  
 Db 672 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATTAATAATATATAT 616

RESULT 13  
 US-09-429-755-169/c  
 Sequence 169, Application US/09429755A

Patent No. US2002011467A1  
 GENERAL INFORMATION:  
 APPLICANT: Fridakis, Tony N.  
 APPLICANT: Smith, John M.  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Mishner, Linda E.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Dillon, Davin C.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 TITRE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
 FILE REFERENCE: 210121.419C6  
 CURRENT APPLICATION NUMBER: US/09/429.755A  
 NUMBER OF SEQ ID NOS: 315  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 169  
 LENGTH: 1265  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-429-755-169

Query Match  
 Best Local Similarity 50.8%; Score 37.8; DB 10; Length 1265;  
 Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

240 AGAATACGAGTACCATTAACCTTGGGTAGTGTGATTTCTGCAATTAACAGTGC 299  
 Db 792 AAATTAACAGGCTTCAGATTACTTGGCTTCATTAATTAATTTTCTTTAAAGAAA 733  
 Oy 300 TTGAACCCATTTCTTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGTT 359  
 Db 732 ATATCAACCATTTGTCAATGACCTGTTTTCAAAGCATTTAAAGAGGTAAACCCCTT 673  
 Oy 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATAT 416  
 Db 672 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATTAATAATATATAT 616

RESULT 14  
 US-09-924-400-290/c  
 Sequence 290, Application US/09924400  
 Patent No. US20020165371A1  
 GENERAL INFORMATION:  
 APPLICANT: Fridakis, Tony N.  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Smith, John M.  
 APPLICANT: Mishner, Linda E.  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yasir A. W.  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Day, Craig H.  
 APPLICANT: Li, Samuel X.  
 APPLICANT: Deng, Ta  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 TITRE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
 FILE REFERENCE: 210121.419C12  
 CURRENT APPLICATION NUMBER: US/09/924.400  
 NUMBER OF SEQ ID NOS: 340  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 290  
 LENGTH: 1646  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-924-400-290

Query Match  
 Best Local Similarity 50.8%; Score 37.8; DB 9; Length 1646;  
 Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

240 AGAATACGAGTACCATTAACCTTGGGTAGTGTGATTTCTGCAATTAACAGTGC 299  
 Oy 300 TTGAACCCATTTCTTACTCTGACCAAGACCATTTAAAGAAATGATTCATCGTT 359  
 Db 732 ATATCAACCATTTGTCAATGACCTGTTTTCAAAGCATTTAAAGAGGTAAACCCCTT 673  
 Oy 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGAGTATCAGAAACATAT 416  
 Db 672 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATTAATAATATATAT 616

Db 1165 AAATTAACAGGCTTCGATTAATCTGGCTTCATATTAATTTCTTTAAAGAAA 1106  
QY 300 TTGAACCCCAATTCCTATATCTGACCAAGACCAATTAAGAAATGATTCATCGGT 359  
Db 1105 ATATCAACCCATGTGATGATGCTGTTTCAAGCATTTAAATAGAGGTAAACCCCT 1046  
QY 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATAT 416  
Db 1045 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATAAAAATTAATATAT 989

## RESULT 15

US-09-810-936-290/c  
; Sequence 290, Application US/09810936  
; Patent No. US20020068285A1  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Smith, John M.  
; APPLICANT: Misher, Linda E.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Day, Craig H.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; FILE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C11  
; CURRENT APPLICATION NUMBER: US/09/810,936  
; CURRENT FILING DATE: 2001-03-16  
; NUMBER OF SEQ. ID NOS: 334  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 290  
; LENGTH: 1646  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-810-936-290

## Query Match

Best Local Similarity 8.0%; Score 37.8; DB 10; Length 1646;  
Matches 90; Conservativity 50.8%; Pred. No. 0.74; Mismatches 87; Indels 0; Gaps 0;

QY 240 AGAATACAGGTACATTAATCTTGGGTAGTGTGTTATCTGCAATTAACAGTGC 299  
Db 1165 AAATTAACAGGCTTCGATTAATCTGGCTTCATATTAATTTCTTTAAAGAAA 1106  
300 TTGAACCCCAATTCCTATATCTGACCAAGACCAATTAAGAAATGATTCATCGGT 359  
Db 1105 ATATCAACCCATGTGATGATGCTGTTTCAAGCATTTAAATAGAGGTAAACCCCT 1046  
QY 360 TTGGCATTAATACAGCAAGAAATCTATGACAGCAAGATGATCAGAAACATAT 416  
Db 1045 TGGAAATTAATACAGAAAGAAATGATTCATTTATGCAATAAAAATTAATATAT 989

Search completed: February 3, 2003, 14:20:03  
Job time : 59 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:01:52 ; Search time 60 Seconds

(without alignments)  
348.672 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 826  
Sequence: 1 AQTYSVALFLGINAFLI.....HMLHSGKCKCHCRCHLS 157

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

1 number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
A. Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	826	100.0	157	22	AA1980.DAT.*
2	638	77.2	188	22	AA1981.DAT.*
3	638	77.2	209	22	AA1982.DAT.*
4	638	77.2	396	21	AA1983.DAT.*
5	638	77.2	396	22	AA1984.DAT.*
6	638	77.2	722	20	AA1985.DAT.*
7	638	77.2	757	20	AA1986.DAT.*
8	632	76.5	176	22	AA1987.DAT.*
9	632	76.5	176	22	AA1988.DAT.*
10	632	76.5	176	22	AA1989.DAT.*

11	632	76.5	176	22	AA1990.DAT.*
12	632	76.5	176	22	AA1991.DAT.*
13	632	76.5	176	22	AA1992.DAT.*
14	632	76.5	176	22	AA1993.DAT.*
15	632	76.5	176	22	AA1994.DAT.*
16	586.5	71.0	140	21	AA1995.DAT.*
17	392.5	47.5	355	22	AA1996.DAT.*
18	392.5	47.5	355	22	AA1997.DAT.*
19	392.5	47.5	491	23	AA1998.DAT.*
20	392.5	47.5	515	23	AA1999.DAT.*
21	392.5	47.5	610	23	AA2000.DAT.*
22	392.5	47.5	713	23	AA2001.DAT.*
23	392.5	47.5	730	23	AA2002.DAT.*
24	392.5	47.5	737	23	AA2003.DAT.*
25	392.5	47.5	754	23	AA2004.DAT.*
26	351	42.5	108	23	AA2005.DAT.*
27	349	42.3	107	23	AA2006.DAT.*
28	349	42.3	109	23	AA2007.DAT.*
29	244	29.5	327	22	AA2008.DAT.*
30	224.5	27.2	334	22	AA2009.DAT.*
31	224.5	27.2	334	22	AA2010.DAT.*
32	212.5	25.7	359	22	AA2011.DAT.*
33	212.5	25.7	359	22	AA2012.DAT.*
34	207.5	25.1	1115	23	AA2013.DAT.*
35	139.5	16.9	620	14	AA2014.DAT.*
36	139.5	16.9	634	14	AA2015.DAT.*
37	139.5	16.9	689	14	AA2016.DAT.*
38	139.5	16.9	690	14	AA2017.DAT.*
39	139.5	16.9	692	11	AA2018.DAT.*
40	139.5	16.9	695	14	AA2019.DAT.*
41	139.5	16.9	695	14	AA2020.DAT.*
42	139.5	16.9	695	14	AA2021.DAT.*
43	139.5	16.9	696	14	AA2022.DAT.*
44	139.5	16.9	696	14	AA2023.DAT.*
45	139.5	16.9	696	14	AA2024.DAT.*

## ALIGNMENTS

RESULT 1  
AAB35407 standard; Protein, 157 AA.  
XX AAB35407;  
AC AAB35407;  
XX 23-MAY-2001 (first entry)  
DT 23-MAY-2001 (first entry)  
XX Human HGR101 G-protein coupled receptor.  
DE Human HGR101 G-protein coupled receptor.  
XX Human; HGR101; G-protein coupled receptor; infection; pain; cancer;  
KW diabetes; obesity; eating disorder; asthma; Parkinson's disease;  
KW hypotension; osteoporosis; myocardial infarction; migraine; allergy;  
XX psychotic disorder; neurological disorder; dyskinesia; vaccine.  
XX Homo sapiens.  
OS Homo sapiens.  
XX WO200114548-A2.  
XX 01-MAR-2001.  
XX 01-MAR-2001.  
XX 09-AUG-2000; 2000MO-EP07723.  
XX 19-AUG-1999; 99EP-0116345.  
XX (MERRE ) MERCK PATENT GMBH.  
XX Duecker K;  
XX WPI; 2001-226617/23.  
XX N-PSDB; AAF28059.  
XX Novel G-coupled protein receptor, HGR101 useful for treating diseases

PT such as microbial infections, cancers, obesity, asthma, diabetes,  
PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer,  
XX allergy -  
XX Claim 1; Page 36; 36pp; English.  
XX  
CC The present invention provides the protein and coding sequences for a  
CC novel human G-protein coupled receptor, designated HGR101. The sequences  
CC are useful in the diagnosis, prevention and treatment of diseases  
CC including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,  
CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
CC vomiting, psychotic and neurological disorders such as anxiety,  
CC schizophrenia, manic depression, depression, delirium, dementia and  
CC severe mental retardation, and dyskinesias including Huntington's disease  
CC and Gilles de la Tourette's syndrome. The present sequence is the HGR101  
CC protein.  
XX  
XX Sequence 157 AA;  
Query Match 100.0%; Score 826; DB 22; Length 157;  
Best Local Similarity 100.0%; Pred. No. 1e-88;  
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 AQTISVAIFLGINLAATITIVFSGSMFYSVHOSATATATINQVKEMILAKRFFIYF 60  
Db 1 AQTISVAIFLGINLAATITIVFSGSMFYSVHOSATATATINQVKEMILAKRFFIYF 60  
CY 61 TDAICWPIPIFAKPLSLLOVEIPGTITSMVWIGSAINSLNLTTLTRPPEKMIHRF 120  
Db 61 TDAICWPIPIFAKPLSLLOVEIPGTITSMVWIGSAINSLNLTTLTRPPEKMIHRF 120  
CY 121 WHNTRKSKMDSKIGIRKMLHSSGKCGHCRCHLS 157  
Db 121 WHNTRKSKMDSKIGIRKMLHSSGKCGHCRCHLS 157  
RESULT 2  
AAM99945  
ID AAM99945 standard; Protein; 188 AA.  
XX  
XX AAM99945;  
XX  
DT 04-JAN-2002 (first entry)  
XX  
DE Human expressed polypeptide SEQ ID NO 69.  
Human; nootropic; neuroprotective; cytosstatic; dermatological; virucide;  
immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
antiparkinsonian; antisticking; antianemic; antiarthritic; cancer;  
KM antineuritic; hepatocytic; cerebroprotective; antiinflammatory;  
KM antiallergic; antidiabetic; antileuker; anticonvulsant; antifungal;  
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KM neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
XX WO200155387-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01110.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 22-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232063.  
PR 14-SEP-2000; 2000US-0232064.  
PR 14-SEP-2000; 2000US-0232065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241825.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246522.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249254.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0251989.  
 PR 06-DEC-2000; 2000US-0251985.  
 PR 06-DEC-2000; 2000US-0251986.  
 PR 06-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

DR WPI; 2001-465573/50.

XX N-PSDB; AA199557.

PT Isolated digestive system associated polypeptide for treating,  
 PT preventing and/or prognosing disorders related to the digestive system  
 PT including digestive system cancers and also for testing and detection  
 PT e.g. diagnosis -

PS Claim 11, SEQ ID NO 69; 509pp + Sequence Listing; English.

CC The invention relates to novel genes (AA199548-AA199604) and proteins  
 CC (AA199936-AA199984) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 188 AA;

Query Match 77.2%; Score 638; DB 22; Length 188;  
 Best Local Similarity 94.8%; Pred. No. 1,4e-66;  
 Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AQTYSVAIFGIMLAAFIIVSYSGSMFYSHQSAITATIRNOVKEMILAKRFFFIIVF 60  
 DB 7 AQTYSVAIFGIMLAAFIIVSYSGSMFYSHQSAITATIRNOVKEMILAKRFFFIIVF 66

QY 61 TDLACWIPFVAKESLLOVEIPGTTSMWVGYSAINALPILYTLTRPFKMIHRF 120  
 DB 67 TDLACWIPFVAKESLLOVEIPGTTSMWVGYSAINALPILYTLTRPFKMIHRF 126

QY 121 MMYNRQRKMSDKG 134  
 DB 127 MMYNRQRKMSDKG 140

RESULT 3  
 ID AAU87243  
 AAU87243 standard; Protein; 209 AA.

AC AAU87243;

DT 05-JUN-2002 (first entry)

XX Novel central nervous system protein #153.

KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
 KW endocrine disorder; diabetes; cancer; leukemia; neovascularisation;  
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
 KW myocardial infarction; wound healing; cell proliferation; skin aging;  
 KW food additive; food preservative; gene therapy.

OS Homo sapiens.

XX WO200155318-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01332.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

[illegible]

FR	20-OCT-2000;	2000US-0241787.
FR	20-OCT-2000;	2000US-0241808.
FR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
FR	08-NOV-2000;	2000US-0246476.
FR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI; 2001-581633/65.	
XX	N-PsDB; ABR43573.	
XX		
PT	New isolated nucleic acid encoding a protein for diagnosing,	
PT	preventing, treating or ameliorating medical conditions and used as	
PT	Food additives or preservatives -	
PS		
XX	Claim 9; SEQ ID No 761; 837pp; English.	
XX		
CC	The invention describes an isolated nucleic acid molecule (I) encoding a	
CC	novel central nervous system protein. (I) and polypeptides (III) encoded	
CC	by (I), are used to treat a medical conditions and in diagnosis of a	
CC	pathological condition. Disorders which are diagnosed or treated include	
CC	autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative	
CC	disorders e.g. neoplasms of the breast or liver, cardiovascular disorders	

Query Match	77.2%;	Score 638;	DB 22;	Length 209;
Fast Local Similarity	94.8%;	Pred. NO. 1.7e-66;		
Matches 127;	Conservative	1;	Mismatches 6;	Indels 0;
				Gaps 0;

RESULT 4  
AAV57286  
ID AAV57286 standard; Protein: 396 AA.

OS	Homo sapiens.
XX	
PN	WO200015793-A2.

PT Hunan G protein coupled protein receptor peptides useful for the  
PT prevention, diagnosis and treatment of cell proliferative, neurological  
PT and immune disorders -  
XX

Query Match	77.2%;	Score 638;	DB 21;	Length 396;
Best Local Similarity	94.8%;	Pred. No. 3.8e-66;		
Matches 127;	Conservative 1;	Mismatches 6;	Indels 0;	Gaps 0;

RESULT 5	
AAE02498	
ID	AAE02498 standard; Protein; 396 AA.

xx  
xx  
os  
xx

**Homo sapiens.**

FT	Domain	42..65	/label= Transmembrane_domain_(1TW)
FT		65..78	/label= Intracellular_domain
FT	Domain	/note= "First IC loop"	
FT		79..103	/label= Transmembrane_domain_(2TW)
FT	Domain	104..124	/label= Extracellular_domain
FT	Domain	/note= "First EC loop"	
FT		125..156	/label= Transmembrane_domain_(3TW)
FT	Domain	157..166	/label= Intracellular_domain
FT		/note= "Second IC loop"	
FT	Domain	167..188	/label= Transmembrane_domain_(4TW)
FT	Domain	189..216	

```
FT /label= Extracellular domain
FT /note= "Second EC loop"
FT 217..241
FT Domain
FT /label= Transmembrane_domain_(5TM)
FT 242..267
FT /label= Intracellular_domain
FT /note= "Third IC loop"
FT 268..290
FT Domain
FT /label= Transmembrane_domain_(6TM)
FT 291..300
FT Domain
FT /label= Extracellular domain
FT /note= "Third EC loop"
FT 301..320
FT Domain
FT /label= Transmembrane_domain_(7TM)
XX WO200131014-A2.
XX
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-US29601.
XX
XX 27-OCT-1999; 99US-0427653.
XX 27-OCT-1999; 99US-0427859.
XX 27-OCT-1999; 99US-0428020.
XX 27-OCT-1999; 99US-0428114.
XX 28-OCT-1999; 99US-0429517.
XX 28-OCT-1999; 99US-0429555.
XX 28-OCT-1999; 99US-0429676.
XX 28-OCT-1999; 99US-0429695.
XX 03-DEC-1999; 99US-0454399.
XX 12-JAN-2000; 2000US-0481794.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Vogel1 G, Wood LS, Merchant K;
XX WPI; 2001-328653/34.
XX N-PSDB; AAD06507.
XX
XX Seven transmembrane receptor polypeptides and polynucleotides, useful
XX for treating neurological or psychiatric disorders, e.g. schizophrenia,
XX as well as for identifying compounds useful for treating schizophrenia
XX
XX Claim 1; Page 15-16; 215pp; English.
XX
XX The invention relates to human G protein-coupled receptor (GPCR) and
XX their corresponding DNA molecules. GPCR is also referred as seven
XX transmembrane receptor. G protein-coupled receptor protein is useful for
XX treating neurological disorder, particularly schizophrenia. GPCR protein
XX is also useful for identifying compounds useful for treating
XX schizophrenia. These compounds are also useful for treating other
XX neurological and psychiatric diseases, e.g. depression, anxiety, bipolar
XX disease, affective disorders, attention deficit hyperactivity disorder/
XX attention deficit disorder, epilepsy, neuritis, neuroasthenia, neuropathy,
XX neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile
XX dementia. The invention also provides genetic screening procedures that
XX entail analyzing a person's genome with respect to GPCR. The vectors are
XX useful for the recombinant production of the GPCR's. The present sequence
XX is human CON222 G protein-coupled receptor (GPCR) protein.
XX
XX Sequence 396 AA;
XX
XX Query Match 77.2%; Score 638; DB 22; Length 396;
XX Best Local Similarity 94.8%; Pred. No. 3.8e-66;
XX Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
Db 275 TDALCWPIPIFVAKPLSLQVEIPGTTISWVIFILPINSALNPILYTLTRPFKEMIRF 334
Oy 121 WHNYRKRKSDSKG 134
Db 335 WHNYRKRKSDSKG 348
XX
XX RESULT 6
XX ID AAY42171
XX AAAY42171 standard; Protein; 722 AA.
XX AC AA442171;
XX XX
XX 20-DEC-1999 (first entry)
XX
XX Human IGR7 short form protein sequence.
XX
XX Human; IGR4; IGR5; IGR7; G-protein coupled receptor; gene therapy;
XX extracellular leucine rich repeat region; mapping; identification.
XX
XX Homo sapiens.
XX OS
XX MO9948921-A1.
XX
XX 30-SEP-1999.
XX
XX 25-MAR-1999; 99MO-US06573.
XX
XX 26-MAR-1998; 98US-0079501.
XX
XX (STRD ) UNIV LEIAND STANFORD JUNIOR.
XX (ORCA ) ORGANON NV.
XX
XX Heueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX WPI; 1999-591074/50.
XX N-PSDB; AA225346.
XX
XX New G-protein coupled receptors, useful for identifying their own
XX ligands -
XX
XX Claim 2; Fig 4; 54pp; English.
XX
XX The present sequence represents the human G-protein coupled receptor
XX having extracellular leucine rich repeat regions, designated IGR7 short
XX form. The IGR4, IGR5 and IGR7 proteins are used to identify ligands for
XX the receptor. The polypeptides and/or polynucleotides are also useful
XX for homologous or related genes, producing compositions that modulate
XX the expression or function of the receptors, gene therapy, mapping
XX functional regions of the receptors, studying associated physiological
XX pathways, in vivo prophylactic and therapeutic purposes, as immunogens
XX for producing antibodies, and for identifying biologically active
XX agents. The polypeptides contain a G-protein coupled seven
XX transmembrane region and a leucine rich repeat extracellular domain.
XX These regions capture and facilitate optimal orientation of its ligand.
XX The proteins are also expressed in diverse tissues.
XX
XX Sequence 722 AA;
XX
XX Query Match 77.2%; Score 638; DB 20; Length 722;
XX Best Local Similarity 94.8%; Pred. No. 8.3e-66;
XX Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```



Db 661 WYNYRORRKSMDSKG 674

RESULT 7

AAV42170  
ID AAV42170 standard; Protein; 757 AA.

AAV42170;

20-DEC-1999 (first entry)

Human LGR7 long form protein sequence.

Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy; extracellular leucine rich repeat region; mapping; identification.

Homo sapiens.

WO9948921-A1.

30-SEP-1999.

25-MAR-1999; 99WO-US06573.

26-MAR-1998; 98US-0079501.

(STRD ) UNIV LELAND STANFORD JUNIOR.  
(ORGA ) ORGANO NV.

Hueh AJW, Hsu SY, Liang S, Van Der Spek PJ;

WPI; 1999-591074/50.

N-PDB; AA25345.

New G-protein coupled receptors, useful for identifying their own ligands -

Claim 2; Fig 3; 54pp; English.

The present sequence represents the human G-protein coupled receptor having extracellular leucine rich repeat regions, designated LGR7 long form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for the receptor. The polypeptides and/or polynucleotides are also useful for homologous or related genes, producing compositions that modulate the expression or function of the receptors, gene therapy, mapping functional regions of the receptors, studying associated physiological pathways, in vivo prophylactic and therapeutic purposes, as immunogens for producing antibodies, and for identifying biologically active agents. The polypeptides contain a G-protein coupled seven transmembrane region and a leucine rich repeat extracellular domain. These regions capture and facilitate optimal orientation of its ligand. The proteins are also expressed in diverse tissues.

Sequence 757 AA;

Query Match 77.2%; Score 638; DB 20; Length 757;

Best Local Similarity 94.8%; Pred. No. 8, 8e-66; Indels 0; Gaps 0;

Matches 127; Conservative 1; Mismatches 6;

1 AQTVAIFLGINIAAFIIIVFSYGSWFSVHOSATATATEIRNWKEMILAKREFFIVF 60

576 AQTVAIFLGINIAAFIIIVFSYGSWFSVHOSATATATEIRNWKEMILAKREFFIVF 635

61 TDALCWIPIPAKRLSLQVEIRGTTISWVVGSAINSANPILYTLTRPKEMIRF 120

636 TDALCWIPIPAKRLSLQVEIRGTTISWVVGSAINSANPILYTLTRPKEMIRF 695

121 WMYRORRKSMDSKG 134

696 WMYRORRKSMDSKG 709

RESULT 8

AAU87545  
ID AAU87545 standard; Protein; 176 AA.

AAU87545;

05-JUN-2002 (first entry)

Novel central nervous system protein #455.

Central nervous system; CNS; autoimmune disease; rheumatoid arthritis; hyperproliferative disorder; neoplasm; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; ischemia; angiodysplasia; nervous system disorder; Alzheimer's disease; AIDS; ocular disorder; acquired immunodeficiency virus; dysphagia; gastrointestinal disorder; adenocarcinoma; reproductive system disorder; testicular feminisation; endocrine disorder; diabetes; cancer; leukemia; neovascularisation; respiratory disorder; renal disorder; kidney failure; blood disorder; myocardial infarction; wound healing; cell proliferation; skin aging; food additive; food preservative; gene therapy.

Homo sapiens.

WO200155318-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US01332.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0214866.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216647.

07-JUL-2000; 2000US-0216880.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

18-AUG-2000; 2000US-0226279.

22-AUG-2000; 2000US-0226281.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

23-AUG-2000; 2000US-0227099.

30-AUG-2000; 2000US-0228924.

01-SEP-2000; 2000US-0228927.

01-SEP-2000; 2000US-0229343.

01-SEP-2000; 2000US-0229344.

01-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229509.

05-SEP-2000; 2000US-0229513.

06-SEP-2000; 2000US-0230437.

06-SEP-2000; 2000US-0230438.

08-SEP-2000; 2000US-0231242.

[illegible]

PR	17-NOV-2000	2000US-02493664
PR	17-NOV-2000	2000US-02493654
PR	17-NOV-2000	2000US-02493659
PR	17-NOV-2000	2000US-02493927
PR	17-NOV-2000	2000US-02493937
PR	17-NOV-2000	2000US-02493900
PR	01-DEC-2000	2000US-02503160
PR	01-DEC-2000	2000US-02503191
PR	05-DEC-2000	2000US-02510330
PR	05-DEC-2000	2000US-02511368
PR	05-DEC-2000	2000US-02516719
PR	06-DEC-2000	2000US-02514729
PR	08-DEC-2000	2000US-02518568
PR	08-DEC-2000	2000US-02518658
PR	08-DEC-2000	2000US-02518693
PR	08-DEC-2000	2000US-02519189
PR	09-DEC-2000	2000US-02513900
PR	11-DEC-2000	2000US-02540977
PR	05-JAN-2001	2001US-02564728

XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-581633/65.  
DR  
DR N-PSDB; ABK43875.  
XX  
XX  
PT New isolated nucleic acid encoding a protein for diagnosing,  
PT Preventing, treating or ameliorating medical conditions and used as  
PT food additives or preservatives -  
XX  
XX Claim 9; SEQ ID No 1063; 837bp; English.

XX The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (II) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Query Match	76.5%	Score 632;	DB 22;	Length 176;
Best Local Similarity	94.0%;	Pred. No. 6.7e-66;		
Matches 126;	Conservative 1;	Mismatches 7;	Indels 0;	Gaps 0

QY	1	AOIYSAVILPLGINLAAPIIIIVYSGSMFYVHOSALTATIEINQVKEMILAKKFFIVE	60
Db	1	AOIYSAVILPLGINLAAPIIIIVYSGSMFYVHOSALTATIEINQVKEMILAKKFFIVE	60
QY	61	TDMACMIPFVAKPIESLLQVEIPGTITSMVYIGSAINALNAPIIYITTTTPFEKMHTRF	120
Db	61	TDMACMIPFVAKPIESLLQVEIPGTITSMVYIGSAINALNAPIIYITTTTPFEKMHTRF	120
QY	121	MHNYRQRKSMDSKG	134
Db	121	MHNYRQRKSMDSKG	134

RESULT 9  
 AAM9972  
 ID AAM9972 standard; Protein; 176 AA.  
 XX  
 AC AAM9972;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX  
 DE Human expressed polypeptide SEQ ID NO 96.  
 XX  
 KM Human; neurotrophic; cytostatic; dermatological; virocidic;  
 KM immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
 KM antiparkinsonian; antischizoid; antineoplastic; cancer;  
 KM antineuritic; hepatotropic; ceredroprotective; antiinflammatory;  
 KM antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 KM antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
 KM neurological disease; infection; neurotrophic; gene therapy; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200155387-A1.  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01310.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 14-AUG-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0225779.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 02-OCT-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM,  
DR WPI; 2001-465573/50.  
DR N-PSDB; AA199584.  
XX  
XX Isolated digestive system associated polypeptide for treating,  
PT preventing and/or prognosing disorders related to the digestive system  
PT including digestive system cancers and also for testing and detection  
PT e.g. diagnosis -  
XX  
XX Claim 11; SEQ ID NO 96; 509pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (AA199548-AA199604) and proteins  
CC (AA199936-AA199984) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX  
XX Sequence 176 AA;  
SQ  
Query Match 76.5%; Score 632; DB 22; Length 176;  
Best Local Similarity 94.0%; Pred. No. 6,7e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 1 AQTYSVAIFLGINLAFTIIVFSGSMFYVHOSATATATIRKQVKEMTLARPPFIYF 60  
DB 1 AQTYSVAIFLGINLAFTIIVFSGSMFYVHOSATATATIRKQVKEMTLARPPFIYF 60  
QY 61 TDAICMIFIFVAKPLSLQVPIPGTITSMVWVIGYSANSLNPLTLTLTRPKEMIHNF 120  
DB 61 TDAICMIFIFVAKPLSLQVPIPGTITSMVWVIGYSANSLNPLTLTLTRPKEMIHNF 120  
QY 121 WHNYRQRKSDSKG 134  
DB 121 WHNYRQRKSDSKG 134  
QY 121 WHNYRQRKSDSKG 134  
DB 121 WHNYRQRKSDSKG 134  
RESULT 10  
ABR04062  
ID ABR04062 standard; Protein; 176 AA.  
XX

AC ABR04062;  
XX  
XX 08-JAN-2002 (first entry)  
DT  
XX  
XX Human musculoskeletal system related polypeptide SEQ ID NO 2009.  
DE  
XX  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system.  
XX  
XX Homo sapiens.  
XX  
XX WO20015367-A1.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01338.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217996.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228927.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR



DE Human cDNA SEQ ID NO: 557.  
XX  
XX Human, gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
XX proliferative disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN MO200154474-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01349.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 15-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190078P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214886P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216880P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225266P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225447P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 14-AUG-2000; 2000US-225759P.  
PR 18-AUG-2000; 2000US-226279P.  
PR 22-AUG-2000; 2000US-226681P.  
PR 22-AUG-2000; 2000US-226868P.  
PR 22-AUG-2000; 2000US-227182P.  
PR 23-AUG-2000; 2000US-227009P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-229287P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 06-SEP-2000; 2000US-230437P.  
PR 06-SEP-2000; 2000US-230438P.  
PR 08-SEP-2000; 2000US-231242P.  
PR 08-SEP-2000; 2000US-231243P.  
PR 08-SEP-2000; 2000US-231244P.  
PR 08-SEP-2000; 2000US-231413P.  
PR 08-SEP-2000; 2000US-231414P.  
PR 08-SEP-2000; 2000US-232080P.  
PR 08-SEP-2000; 2000US-232081P.  
PR 12-SEP-2000; 2000US-232968P.  
PR 14-SEP-2000; 2000US-232397P.  
PR 14-SEP-2000; 2000US-232398P.  
PR 14-SEP-2000; 2000US-232399P.  
PR 14-SEP-2000; 2000US-232400P.  
PR 14-SEP-2000; 2000US-232401P.  
PR 14-SEP-2000; 2000US-233063P.  
PR  
PR 14-SEP-2000; 2000US-233064P.  
PR 14-SEP-2000; 2000US-233065P.  
PR 21-SEP-2000; 2000US-234223P.  
PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234998P.  
PR 25-SEP-2000; 2000US-234999P.  
PR 26-SEP-2000; 2000US-235834P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 27-SEP-2000; 2000US-235835P.  
PR 29-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-239935P.  
PR 13-OCT-2000; 2000US-239937P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241221P.  
PR 20-OCT-2000; 2000US-241785P.  
PR 20-OCT-2000; 2000US-241786P.  
PR 20-OCT-2000; 2000US-241787P.  
PR 20-OCT-2000; 2000US-241808P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 20-OCT-2000; 2000US-241826P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 08-NOV-2000; 2000US-246474P.  
PR 08-NOV-2000; 2000US-246475P.  
PR 08-NOV-2000; 2000US-246476P.  
PR 08-NOV-2000; 2000US-246477P.  
PR 08-NOV-2000; 2000US-246478P.  
PR 08-NOV-2000; 2000US-246523P.  
PR 08-NOV-2000; 2000US-246524P.  
PR 08-NOV-2000; 2000US-246525P.  
PR 08-NOV-2000; 2000US-246526P.  
PR 08-NOV-2000; 2000US-246527P.  
PR 08-NOV-2000; 2000US-246528P.  
PR 08-NOV-2000; 2000US-246532P.  
PR 08-NOV-2000; 2000US-246609P.  
PR 08-NOV-2000; 2000US-246610P.  
PR 08-NOV-2000; 2000US-246611P.  
PR 08-NOV-2000; 2000US-246613P.  
PR 17-NOV-2000; 2000US-249207P.  
PR 17-NOV-2000; 2000US-249208P.  
PR 17-NOV-2000; 2000US-249209P.  
PR 17-NOV-2000; 2000US-249210P.  
PR 17-NOV-2000; 2000US-249211P.  
PR 17-NOV-2000; 2000US-249212P.  
PR 17-NOV-2000; 2000US-249213P.  
PR 17-NOV-2000; 2000US-249214P.  
PR 17-NOV-2000; 2000US-249215P.  
PR 17-NOV-2000; 2000US-249216P.  
PR 17-NOV-2000; 2000US-249217P.  
PR 17-NOV-2000; 2000US-249218P.  
PR 17-NOV-2000; 2000US-249244P.  
PR 17-NOV-2000; 2000US-249245P.  
PR 17-NOV-2000; 2000US-249246P.  
PR 17-NOV-2000; 2000US-249265P.  
PR 17-NOV-2000; 2000US-249297P.  
PR 17-NOV-2000; 2000US-249299P.  
PR 17-NOV-2000; 2000US-249300P.  
PR 01-DEC-2000; 2000US-250160P.  
PR 05-DEC-2000; 2000US-250391P.  
PR 05-DEC-2000; 2000US-251030P.  
PR 05-DEC-2000; 2000US-251988P.  
PR 05-DEC-2000; 2000US-256719P.  
PR 06-DEC-2000; 2000US-251479P.  
PR 08-DEC-2000; 2000US-251856P.  
PR 08-DEC-2000; 2000US-251868P.

PR 08-DEC-2000; 2000US-251869P.  
PR 08-DEC-2000; 2000US-251989P.  
PR 08-DEC-2000; 2000US-251990P.  
PR 11-DEC-2000; 2000US-254097P.  
PR 05-JAN-2001; 2001US-259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-476161/51.  
DR N-PSDB; ABA06471.  
XX  
PT Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
XX condition -  
XX  
PS Claim 11; SEQ ID NO: 557; 859pp + Sequence Listing; English.  
XX  
CC The present invention provides human cDNAs, proteins and related genomic  
CC DNAs. These can be used in the treatment of neural, immune system,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal and proliferative disorders and inflammation. The present sequence  
XX is a protein of the invention.  
SQ Sequence 176 AA;  
Query Match 76.5%; Score 632; DB 22; Length 176;  
Best Local Similarity 94.0%; Pred. No. 6,7e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 1 AOTSVAIFGIMAAPIIVFSYSGMPYSYHOSAITETIRNOVKEMILAKRPFIIV 60  
DB 1 AOTSVAIFGIMAAPIIVFSYSGMPYSYHOSAITETIRNOVKEMILAKRPFIIV 60  
QY 61 TDALCWPPIPVAKFSLQVEIPGTTISWVIGSAINSLPILYLTTRPKEMIHNF 120  
DB 61 TDALCWPPIPVAKFSLQVEIPGTTISWVIGSAINSLPILYLTTRPKEMIHNF 120  
QY 121 WNTYRQRKSMDSKG 134  
DB 121 WNTYRQRKSMDSKG 134  
RESULT 12  
AAU18108  
ID AAU18108 standard; Protein; 176 AA.  
QY  
AAU18108;  
DT 21-NOV-2001 (first entry)  
XX  
DE Novel human uterine motility-association polypeptide #15.  
XX  
KM Human, uterine motility-association disorder; uterus; pregnancy;  
KM labour; menstrual cycle; gene therapy.  
XX  
OS Homo sapiens.  
XX  
XX WO200155201-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01317.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232081.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235884.  
PR 27-SEP-2000; 2000US-0235884.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246612.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256179.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-488777/53.  
XX N-PSDB; AAS28950.  
XX  
XX Isolated polypeptide and nucleic acid molecules for treating,  
XX preventing and/or prognosing disorders related to uterine motility  
XX e.g. disorders associated with pregnancy and the menstrual cycle -  
XX  
XX Claim 11; SEQ ID No 84; 524pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
XX uterine motility-association polypeptides, and cDNA (AAS28936-AAS28994)  
XX and genomic sequences encoding for these polypeptides.  
XX The sequences of the invention are useful in the diagnosis,

CC treatment, prevention and/or prognosis of diseases associated  
CC with uterine motility such as pregnancy and labour, and menstrual  
CC disorders. The polynucleotide sequences of the invention are also  
CC useful in gene therapy. AAU18094-AAU18152 represent novel human  
CC uterine motility-association polypeptides.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 176 AA;  
SQ  
Query Match 76.5%; Score 632; DB 22; Length 176;  
Best Local Similarity 94.0%; Pred. No. 6,7e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
OY 1 AAIYSVAIFPLGINAAFIIVPSYSGMFYSVHQSATATETIRNOVKEMILAKRFFFIIV 60  
DB 1 AAIYSVAIFPLGINAAFIIVPSYSGMFYSVHQSATATETIRNOVKEMILAKRFFFIIV 60  
OY 61 TDALCWIPIPVAKPLSLGVEIPGTTISWVVTGSAINSLNPLVLTITRPFKMIRF 120  
DB 61 TDALCWIPIPVAKPLSLGVEIPGTTISWVVTGSAINSLNPLVLTITRPFKMIRF 120  
OY 121 WHNYRKRKSMDSKG 134  
DB 121 WHNYRKRKSMDSKG 134  
RESULT 13  
AAU18344  
ID AAU18344 standard; Protein; 176 AA.  
XX  
XX AAU18344;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
DT  
XX  
XX Human endocrine polypeptide SEQ ID No 299.  
DE  
XX  
XX Endocrine protein; human; mouse; rabbit; goat; horse; food additive;  
KW cat; dog; chicken; sheep; immunosuppressive; antidiabetic; vasotropic;  
KW antihemnetic; antiproliferative; cytostatic; cardiatic; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; viucide; fungicide; cancer;  
KW ophthalmological; vunerary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-infertility.  
OS  
XX Homo sapiens.  
XX  
XX WO200155164-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01308.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0188874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.



PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 14-AUG-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 18-AUG-2000; 2000US-0225759.  
 PR 22-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0232081.  
 PR 14-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 PT WPI, 2001-451936/48.  
 XX N-PSDB; AAS29573.  
 DR  
 XX  
 PT Isolated polypeptide for treating, preventing and/ or prognosing  
 PT disorders of the endocrine system such as reproductive disorders,  
 PT endocrine cancers and also for testing and detection e.g. diagnosis -  
 PS  
 PS Claim 11; SEQ ID No 299; 604pp; English.  
 XX  
 CC Sequences AAU18282-AAU18507 represent endocrine polypeptides of the  
 CC invention. Endocrine polypeptides and their associated polynucleotides  
 CC are useful in the diagnosis, treatment and prevention of various types of  
 CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. A pathological condition can be determined by  
 CC determining the presence or absence of a mutation in an endocrine  
 CC polynucleotide. The treatable disorders include autoimmune diseases such  
 CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
 CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
 CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
 CC disorders such as Alzheimer's disease, infections caused by bacteria.

CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
CC disorders such as premature labour and infertility, gastrointestinal  
CC disorders such as Crohn's disease, renal disorders such as  
CC glomerulonephritis and respiratory disorders such as asthma. The  
CC polypeptides can also be used to aid wound healing, to prevent skin aging  
CC due to sunburn, to maintain organs before transplantation, to regenerate  
CC tissues and in chemotaxis. The polypeptides can also be used as a food  
CC additive or preservative to increase or decrease storage capabilities.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at fcp.wipo.int/pub/published\_pct\_sequences.

Query Match 76.5%; Score 632; DB 22; Length 176;  
Best Local Similarity 94.0%; Pred. No. 6,7e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AQTSAVAFPGINAAFFITVFSYSGMFGVHSGAATRTIRNOVKEMILAKRFFIYF 60  
1 AQTSAVAFPGINAAFFITVFSYSGMFGVHSGAATRTIRNOVKEMILAKRFFIYF 60

Db 61 TDALCWIPFVAKPLSLQVEIPGTITSWVYIGYSAINSLNPILYTLTRPKEMIHHP 120  
61 TDALCWIPFVAKPLSLQVEIPGTITSWVYIGYSAINSLNPILYTLTRPKEMIHHP 120

QY 121 WHNYRQRKSMDSKG 134  
121 WHNYRQRKSMDSKG 134

Db 121 WYNYRQRKSMDSKG 134  
121 WYNYRQRKSMDSKG 134

RESULT 14  
AAU18666  
ID AAU18666 standard; Protein: 176 AA.  
XX AAU18666;  
AC AAU18666;  
XX 21-NOV-2001 (first entry)  
DT 21-NOV-2001 (first entry)  
XX  
DE Renal and cardiovascular-associated protein, Seq ID 105.  
XX

Human; antiinflammatory; neuroprotective; immunomodulator; vulnerrary;  
KM cardiovascular; cytoskeletal; nephrotoxic; antihaemic; nephritis;  
KM immunosuppressive; kidney disorder; renal failure; hypertension;  
KM cardiovascular disorder; myocardial infarction; blood disorder; anaemia;  
KM blood coagulation disorder; electrolyte imbalance disorder; cancer;  
KM hyponatremia; hyperkalemia; neoplastic disorder; nephroma;  
KM autoimmune disease; inflammatory disease; reproductive system disorder;  
KM endocrine disorder; neural activity; neurological disorder;  
KM wound healing; respiratory disorder.

XX Homo sapiens.  
XX  
XX WO20015328-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01359.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234897.  
PR 25-SEP-2000; 2000US-0234898.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235634.  
PR 27-SEP-2000; 2000US-0235636.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 06-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246539.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249267.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250161.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251031.  
 PR 05-DEC-2000; 2000US-0251038.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-488787/53.  
 XX N-PSDB; AAS30187.  
 PT New polynucleotides and polypeptides, useful for diagnosing, treating,  
 PT preventing or prognosing e.g. kidney, cardiovascular, blood,  
 PT electrolyte imbalance or neoplastic disorders, autoimmune diseases,  
 PT cancers -  
 XX  
 PS Claim 1; SEQ ID No 105; 506pp; English.  
 PS  
 CC The invention relates to novel nucleic acids and polypeptides useful for  
 CC diagnosing, treating, preventing and/or prognosing disorders related to  
 CC these polypeptides. The polynucleotides are especially useful in the  
 CC diagnosis, prognosis, prevention and/or treatment of diseases which  
 CC include kidney disorders (e.g. renal failure or nephritis),  
 CC cardiovascular disorders (e.g. hypertension or myocardial infarction),  
 CC blood disorders (e.g. anaemia or blood coagulation disorders),  
 CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),  
 CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune  
 CC diseases, cancers, inflammatory diseases, reproductive system  
 CC disorders, endocrine disorders, neural activity and neurological

CC disorders, wound healing and respiratory disorders. AAU18644-AAU18715  
 CC represent the novel human renal and cardiovascular-associated amino  
 CC acid sequences of the invention. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at:  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 176 AA;  
 Query Match 76.5%; Score 632; DB 22; Length 176;  
 Best Local Similarity 94.0%; Pred. No. 6.7e-66;  
 Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 AQTYSVAIFLGNLAFFIIIVSYSGMFYSVQSAITATIRINQVKEMTLAKRFFPIVF 60  
 DB 1 AQTYSVAIFLGNLAFFIIIVSYSGMFYSVQSAITATIRINQVKEMTLAKRFFPIVF 60  
 QY 61 TDALCWIPFIFFVAKPLSLQVEIPGTTSMVVTGYGASINSLNPILYTLTTRPFKEMHIF 120  
 DB 61 TDALCWIPFIFFVAKPLSLQVEIPGTTSMVVTGYGASINSLNPILYTLTTRPFKEMHIF 120  
 QY 121 WNYRQRKSMDSKG 134  
 DB 121 WNYRQRKSMDSKG 134  
 RESULT 15  
 AAU21646  
 ID AAU21646 standard; Protein; 176 AA.  
 XX  
 AC AAU21646;  
 XX  
 DT 06-DEC-2001 (first entry)  
 XX  
 DE Novel human neoplastic disease associated polypeptide #79.  
 XX  
 KW Human; neoplastic disease associated polypeptide; cancer;  
 KW hyperproliferative disorder; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20015163-A1.  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01358.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.

[illegible]

PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249219.
PR	17-NOV-2000;	2000US-0249220.
PR	17-NOV-2000;	2000US-0249221.
PR	17-NOV-2000;	2000US-0249222.
PR	17-NOV-2000;	2000US-0249223.
PR	17-NOV-2000;	2000US-0249224.
PR	17-NOV-2000;	2000US-0249225.
PR	17-NOV-2000;	2000US-0249226.
PR	17-NOV-2000;	2000US-0249227.
PR	17-NOV-2000;	2000US-0249228.
PR	17-NOV-2000;	2000US-0249229.
PR	17-NOV-2000;	2000US-0249230.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251858.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0255978.
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
XX	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI;	2001-465558/50.
DR	N-PSDB;	AAS34845.
XX		
PT	Novel polypeptides and polynucleotides useful as diagnostic reagents to	
PT	diagnose diseases or disorders associated with aberrant expression or	
PT	activity of polypeptides, and for treating cancers, rheumatoid	
PT	arthritis -	
XX		
PS	Claim 11; SEQ ID No 373; 687bp; English.	
XX		
CC	The present invention relates to the isolation of novel human neoplastic	
CC	disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA	
CC	sequences encoding for these polypeptides. The sequences of the	
CC	invention are useful in the diagnosis, treatment, prevention and/or	
CC	prognosis of disorders involving neoplastic disease such as	
CC	hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder	
CC	cancer, brain stem glioma, adult liver cancer, childhood cerebellar	
CC	astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may	
CC	also be useful for treating other disorders such as neural disorders,	
CC	immune system disorders, muscular disorders, reproductive disorders,	
CC	gastrointestinal disorders, pulmonary disorders, cardiovascular disorders	
CC	and renal disorders. The polynucleotide sequences of the invention are	
CC	also useful in gene therapy. AAU21568-AAU21851 represent the novel human	
CC	neoplastic disease associated polypeptides of the invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
CC		
XX		

**SQ Sequence 176 AA;**

Query Match	Score	DB	Length
76.5%	632	22	176

Best Local Similarity 94.0%; Pred. No. 6.7e-66;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

[illegible]

AOIYSVAIFLGINIAAFIIVFSYSGMFYSVHQASATATEIRNQVKKEMILAKRFFIVF 60

Qy 61 TDALCWIPIPIAKPLSLQLVEIPGTITSWVIGISAINSLNPILYTLTRPPKEMIHRF 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 61 TDAICWIPFIVVKPFLSLQLVEIPGTITSWWVIFLPINSALNPIYTLTRPPKEMIHRF 120

61 TDALCMIPFVVKELSLQVEIPGTITSWVIFILPINSALNPILYTLTRPPKEMIHFR 120

QY	121	WHNYRQRKSMDSKG	134
----	-----	----------------	-----

Db 121 WYNYRQRKSMDSXG 134

Search completed: February 3, 2003, 14:21:13  
 Elapsed time : 61 secs

```
time : 61 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:48:12 ; Search time 37 Seconds  
(without alignments)  
407.922 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 826  
Sequence: 1 AQLYSVALFLGINAFLII.....HMLHSSGKCGKCHRCRCHUS 157

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:1\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207.5	25.1	1115	2 S40241	G protein-coupled
2	143.5	17.4	925	2 JCI493	foliitropin recept
3	143	17.3	925	2 JCI493	G protein-coupled
4	141.5	17.1	695	1 JN0898	foliitropin recept
5	140.5	17.0	686	2 A41344	lutropin-receptor
6	139.5	16.9	692	2 A34548	foliitropin recept
7	139.5	16.9	692	2 A34548	foliitropin recept
8	136.5	16.5	700	2 JCI493	foliitropin recept
9	134	16.2	700	2 JCI493	lutropin-receptor
10	134	16.2	700	2 A49744	lutropin-receptor
11	134	16.2	700	2 A49744	lutropin-receptor
12	134	16.2	700	2 JCI493	thyroid stimulat
13	132	16.0	783	2 JCI493	thyroid stimulat
14	131.5	15.9	694	2 JCI493	foliitropin recept
15	131	15.9	699	1 QRHUT	lutropin-choriogon
16	128	15.5	320	2 S17177	probable G protein
17	128	15.5	337	2 S68678	adenosine receptor
18	127	15.4	764	2 JCI493	thyroid stimulat
19	125.5	15.2	695	1 QRHUT	foliitropin recept
20	124	15.0	696	2 JCI493	foliitropin recept
21	124	15.0	764	2 A35856	thyrotropin recept
22	122	14.8	764	2 A40077	thyrotropin recept
23	121	14.6	764	2 A48882	thyrotropin recept
24	119.5	14.5	332	2 A42171	A2-adenosine recep
25	119.5	14.5	378	2 JCI493	G protein-coupled
26	119.5	14.5	417	2 E30341	alpha-1B-adrenergic
27	119.5	14.5	517	2 A45121	alpha-1B adrenergic
28	118	14.3	332	2 I48933	adenosine receptor
29	117.5	14.2	515	2 A40491	alpha-1-adrenergic

30	117.5	14.2	764	1 QRHURH	thyrotropin recept
31	117	14.2	326	2 I48096	A1 adenosine recep
32	117	14.2	326	2 A31444	adenosine receptor
33	117	14.2	332	2 JCI493	adenosine receptor
34	115	13.9	326	2 A53005	adenosine receptor
35	114	13.8	466	2 T26458	hypothetical prote
36	113.5	13.7	320	2 A46152	A3 adenosine recep
37	113.5	13.7	392	2 S65693	opioid receptor mu
38	113.5	13.7	400	2 S65693	mu opiate receptor
39	113	13.7	326	2 C30341	G protein-coupled
40	113	13.7	444	2 T27866	hypothetical prote
41	112	13.6	328	2 JN0675	adenosine receptor
42	111	13.4	185	2 I48931	adenosine receptor
43	111	13.4	326	2 A40376	adenosine receptor
44	111	13.4	386	1 S34043	oxytocin receptor
45	110.5	13.4	324	2 S55302	A(1) adenosine rec

## ALIGNMENTS

### RESULT 1

S40241  
G protein-coupled receptor - great pond snail  
C:Species: Lymnaea stagnalis (great pond snail)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Jun-2001  
C:Accession: S40241

R:Rensen, C.P.; Kesteren, E.R.; Planta, R.U.; Cox, K.; Burke, J.F.; Heerikhuizen, H.; Vre submitted to the EMBL Data Library, June 1993  
A:Description: A G protein-coupled receptor with LDL-binding motifs suggests a role for ]  
A:Reference number: S40241  
A:Accession: S40241  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1115 <1EN>

A:Cross-References: EMBL:Z23104, NID:q438128, PID:q438129  
C:Superfamily: great pond snail LDL receptor-related G protein-coupled receptor; LDL rec  
C:Keywords: G protein-coupled receptor; transmembrane protein  
F:38-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:79-113/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:118-153/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F:158-194/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F:195-230/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F:233-267/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F:274-316/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F:322-361/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F:367-401/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F:406-440/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
F:446-483/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
F:488-523/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
F:584-607/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:608-631/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:632-655/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:656-679/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:704-727/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:774-797/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

Query Match 25.1% Score 207.5; DB 2; Length 1115;  
Best Local Similarity 33.1%; Pred. No. 6.9e-13;  
Matches 51; Conservative 25; Mismatches 45; Indels 33; Gaps 5;

QY	4	YSVALFLINIAAFLIIVFSYGSWFSV--HOSATATEINQYKEMILAKRFFIVF	60
DB	938	YVGVFIILNLSFVLASSTLMNF-SVAKRTSAVTRTESKN-----DNAMRRTILVM	992
QY	61	TDALCWIDIFVAKPLSLQVEIPETTSWVIGVSAINSALNPILYTLTTRPF-----	113
DB	993	TDPCWVPIIIVGFVSISAGRADQVAVIAVFLPINSATPVIYTLTSTAFILGNVRR	1052
QY	114	-----KEMIRH-----MENVORKS	129
DB	1053	ANRFKSTHSFTGDTKHSYVDGTHSYCEKKS	1086

## RESULT 2

JC1493

Follicle-stimulating hormone receptor

N:Alternate names: follicle stimulating hormone receptor

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jan-2000

A:Accession: JC1493; 147080

R:Khan, H.; Yarney, T.A.; Saitem, M.R.

Biochem. Biophys. Res. Commun. 190, 888-894, 1993

A:Title: Cloning of alternatively spliced mRNA transcripts coding for variants of ovine testis

A:Reference number: JC1493; MUID:9316195; PMID:8439338

A:Accession: JC1493

A:Molecule type: mRNA

A:Residues: 1-695 &lt;KHA&gt;

A:Experimental source: testis

R:Yarney, T.A.; Saitem, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seldah, N.G.

Mol. Cell. Endocrinol. 93, 219-228, 1993

A:Title: Molecular cloning and expression of the ovine testicular follicle stimulating hormone

A:Reference number: 147080; MUID:93351750; PMID:8394255

A:Accession: 147080

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-695 &lt;VAR&gt;

A:Cross-references: GB:L07302; NID:g165884; PIDN:AAA1525.1; PID:g165885

C:Genetics:

A:Gene: FSH-R

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR&gt;

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:191-199/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:652-673/Domain: transmembrane #status predicted <TM4>  
 F:692-712/Domain: transmembrane #status predicted <TM5>  
 F:740-763/Domain: transmembrane #status predicted <TM6>  
 F:775-795/Domain: transmembrane #status predicted <TM7>  
 F:61-152,212,435,491,495/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:553,645,855,859,868,894,900/Binding site: phosphate (Ser) (covalent) #status predicted  
 F:649,855/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 17.4% Score 143.5; DB 2; Length 925;  
 Best Local Similarity 33.3%; Pred. No. 2.1e-06;  
 Matches 37; Conservative 21; Mismatches 47; Indels 6; Gaps 3;

QY 6 VAIFGILNLAFFITVSGSMFVSQSAITATEIRNQVKK-EMILAKRFFVFTDAL 64  
 DB 694 VAFLLVNGASFISVMYLSRMLYV---VSGDEGAPKENDSVAKRMALVFTDML 749

QY 65 CWIRP-FVAKPLSLQVEIPGTTISWVIGYSANALPILYTLTPPK 114  
 DB 750 CWAPIAFGLAARGQTLTVTQSKILVFPFINSICNPFYAFPTKAFK 800

## RESULT 4

JN0898

Follicle-stimulating hormone precursor - crab-eating macaque

N:Alternate names: follicle-stimulating hormone receptor (FSHR)

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

A:Accession: JN0898; S36452

R:Gronow, J.; Sharma, R.S.; Nieschlag, E.

Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993

A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of the

A:Reference number: JN0898; MUID:94071854; PMID:7504463

A:Accession: JN0898

A:Molecule type: mRNA

A:Residues: 1-695 &lt;GEO&gt;

A:Cross-references: EMBL:X74454; NID:g396801; PIDN:CAA52463.1; PID:g396802

A&gt;Note: the authors translated the codon AGT for residue 488 as Arg

C:Function:

A:Description: receptor that mediates the biochemical effects of follicle-stimulating hormone

A:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein; h

F:1-11/Domain: signal sequence #status predicted &lt;SIG&gt;

F:18-695/Product: follicle-stimulating hormone receptor #status predicted &lt;PR&gt;

F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR2&gt;

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR3&gt;

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR4&gt;

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR5&gt;

F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR6&gt;

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology &lt;LRR7&gt;

F:367-387/Domain: transmembrane #status predicted &lt;TM1&gt;

F:399-421/Domain: transmembrane #status predicted &lt;TM2&gt;

F:444-465/Domain: transmembrane #status predicted &lt;TM3&gt;

F:486-507/Domain: transmembrane #status predicted &lt;TM4&gt;

F:529-550/Domain: transmembrane #status predicted &lt;TM5&gt;

F:574-597/Domain: transmembrane #status predicted &lt;TM6&gt;

F:609-630/Domain: transmembrane #status predicted &lt;TM7&gt;

F:191-199,293,318/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted

F:555/Binding site: phosphate (Thr) (covalent) #status predicted



Db 640 LLSKFGCYEMOQIYRTETSTSTA-----HNSHPNGHSSAH 676

## RESULT 5

A41344  
Intropin-choriogonadotropin receptor precursor (splice form A) - pig  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 21-Jan-2000  
C/Accession: A41344  
R/Locus: H.; Mistrahl, M.; Atger, M.; Salese, R.; Vu Hai-Luu Thi, M.T.; Jolivet, A.;  
Science 245, 525-528, 1989  
A/Title: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking trans-  
A/Reference number: A41344; MUID:89332517; PMID:2502844  
A/Accession: A41344  
A/Molecule type: mRNA  
A/Residues: 1-696 <R00>  
A/Cross-references: GB:M29525; NID:g164528; PIDN:AAA1062.1; PID:g164529  
C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane pro-  
F/27/Domain: signal sequence #status predicted <SIG>  
F/28-696/Product: Intropin-choriogonadotropic hormone receptor #status predicted <MAT>  
F/99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 17.0%; Score 140.5; DB 2; Length 696;  
Best Local Similarity 27.3%; Pred. No. 2.8e-06;  
Matches 41; Conservative 26; Mismatches 55; Indels 29; Gaps 4;

Qy 1 AOISVAIFLGINLAAPFIIVSYGSMFVSQSAITATEIRNOVKEMILAKRFFIIVF 60  
Db 524 SOLYVWML-LVNVAVAFVVICGCTTHIYLVNPNITSS-----SSDTKIARKMAMLIIF 576  
Qy 61 TDALCWIPF-FVAKPLSLQVEIPGTTTSWVIVGYSAINSALNPILYTLTPRPFKE---- 119  
Db 577 TDFLCMAPIFFPAISASLKVPLITVNSKVLVLFYPVNSCANPFLVAIFTKNFRDFFI 631  
Qy 120 FWHNTQRKSKMSDKIRKEMHSHSGKCGH 150  
Db 632 -----RDFILLKSGCKH 646

## RESULT 6

A34548  
Follicotropin receptor precursor - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 13-Aug-1999  
C/Accession: A34548; A41729  
R/Locus: H.; Braun, T.; Nikolic, K.; Segaloff, D.L.; Seeburg, P.H.  
Mol. Endocrinol. 4, 525-530, 1990  
A/Title: The testicular receptor for follicle stimulating hormone: structure and function  
A/Reference number: A34548; MUID:91125358; PMID:2126341  
A/Accession: A34548  
A/Molecule type: mRNA  
A/Residues: 1-692 <SPR>  
A/Cross-references: GB:L02842; NID:g204183; PIDN:AAA41175.1; PID:g204184  
R/Heckert, D.L.; Daley, I.J.; Griswold, M.D.  
Mol. Endocrinol. 6, 70-80, 1992  
A/Title: Structural organization of the follicle-stimulating hormone receptor gene.  
A/Reference number: A41729; MUID:92149579; PMID:11738373  
A/Accession: A41729  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-692 <HEC>  
A/Cross-references: GB:S81196; NID:g245344; PIDN:AA821415.1; PID:g245345  
A/Note: sequence inconsistent with the nucleotide translation  
A/Note: sequence extracted from NCBI backbone (NCBIN:81117, NCBIN:81119, NCBIN:81121, NC  
Mol. Endocrinol. 9, 159-170, 1995  
A/Title: Identification of the sites of N-linked glycosylation on the follicle-stimulating  
A/Reference number: A57562; MUID:95295729; PMID:7776966  
A/Contents: annotation; glycosylation sites  
C/Function:

A/Description: receptor that mediates the biochemical effects of follicotropin  
C/Suprafamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat ho  
C/Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone recei  
F/1-15/Domain: signal sequence #status predicted <SIG>  
F/16-692/Product: follicotropin receptor #status predicted <MAT>  
F/16-692/Domain: extracellular hormone binding #status predicted <EBH>  
F/16-366/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F/16-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F/16-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F/16-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F/121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F/146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F/172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F/194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F/367-387/Domain: transmembrane #status predicted <TM1>  
F/398-421/Domain: transmembrane #status predicted <TM2>  
F/443-465/Domain: transmembrane #status predicted <TM3>  
F/486-508/Domain: transmembrane #status predicted <TM4>  
F/529-550/Domain: transmembrane #status predicted <TM5>  
F/574-597/Domain: transmembrane #status predicted <TM6>  
F/609-630/Domain: transmembrane #status predicted <TM7>  
F/191-199,293/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/554/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
F/595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 16.9%; Score 139.5; DB 2; Length 692;

Best Local Similarity 27.3%; Pred. No. 3.6e-06;  
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;

Qy 1 AOISVAIFLGINLAAPFIIVSYGSMFVSQSAITATEIRNOVKEMILAKRFFIIVF 60  
Db 526 SOLYVWML-LVNVAVAFVVICGCTTHIYLVNPNITSS-----SSDTKIARKMAMLIIF 578  
Qy 61 TDALCWIPF-FVAKPLSLQVEIPGTTTSWVIVGYSAINSALNPILYTLTPRPFKE---- 115  
Db 579 TDFLCMAPIFFPAISASLKVPLITVNSKVLVLFYPVNSCANPFLVAIFTKNFRDFFI 638  
Qy 116 MIHRF-----WHNTQRKSKMSDKIRK-----MHSS 144  
Db 639 LLSKFGCYEMOQIYRTETSTSTAHPHARKSHCSSAPRVTSYLVLPNLHSS 690

## RESULT 7

I45896  
Follicle stimulating hormone receptor - bovine  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 15-Oct-1996 #sequence\_revision 15-Oct-1996 #text\_change 21-Jan-2000  
C/Accession: I45896  
R/Locus: H.; Lambert, A.; Saunade, J.; Silversides, D.W.; Lussier, J.G.  
Mol. Reprod. Dev. 39, 127-135, 1994  
A/Title: Structure of the bovine follicle-stimulating hormone receptor complementary DNA  
A/Reference number: I45896; MUID:95127199; PMID:7826612  
A/Accession: I45896  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-695 <RDV>  
A/Cross-references: GB:L2319; NID:g404671; PIDN:AAC37324.1; PID:g404672  
C/Genetics:  
A/Suprafamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat ho  
F/71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 16.9%; Score 139.5; DB 2; Length 695;  
Best Local Similarity 26.5%; Pred. No. 3.6e-06;  
Matches 43; Conservative 33; Mismatches 59; Indels 27; Gaps 6;

Qy 1 AOISVAIFLGINLAAPFIIVSYGSMFVSQSAITATEIRNOVKEMILAKRFFIIVF 60  
Db 527 SOLYVWML-LVNVAVAFVVICGCTTHIYLVNPNITSS-----SSDTKIARKMAMLIIF 579  
Qy 61 TDALCWIPF-FVAKPLSLQVEIPGTTTSWVIVGYSAINSALNPILYTLTPRPFKE---- 115  
Db 580 TDFLCMAPIFFPAISASLKVPLITVNSKVLVLFYPVNSCANPFLVAIFTKNFRDFFI 639

```

Oy 116 MHRFW-----NHYRQKSDSKGIKMLHSHSGKCHC 151
Db 640 LSKFGCYEVOAQTYRSTSTSTA-----HNHFPRNGHC 672

RESULT 8
JC4301
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C:Accession: JC4301
R:Gemy, J.J., Labhdi-Mansais, Y., Yerle, M., Bozon, V., Couture, L., Pajot, E., Grebert,
F. 163, 257-261, 1995
A:title: The porcine follicotropin receptor: cDNA cloning, functional expression and chromo
A:Reference number: JC4301; MUID:96011644; PMID:7590277
A:Accession: JC4301
A:Molecule type: mRNA
A:Residues: 1-694 <REM>
C:Cross-references: GB:131966
C:Comment: This receptor belongs to the family of the G-protein coupled receptors. It ha
ematogenesis in male and oogenesis in female.
C:Genetics:
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-35/Domain: follicle-stimulating hormone binding #status predicted <HDB>
F:70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:166-388/Domain: transmembrane #status predicted <TM1>
F:398-420/Domain: transmembrane #status predicted <TM2>
F:443-464/Domain: transmembrane #status predicted <TM3>
F:485-507/Domain: transmembrane #status predicted <TM4>
F:528-549/Domain: transmembrane #status predicted <TM5>
F:573-596/Domain: transmembrane #status predicted <TM6>
F:608-629/Domain: transmembrane #status predicted <TM7>

Query Match 16.5%; Score 136.5; DB 2; Length 694;
Best Local Similarity 26.8%; Pred. No. 7.2e-06;
Matches 42; Conservative 36; Mismatches 62; Indels 17; Gaps 6;

Oy 1 AQIYVAIFGLINLAFFITVPSGMSFYVHQSATITAEIRNQVKEMILAKFFFLV 60
Db 526 SOLYVSL-LVNLVLAFAVITCGCIHIITLVNPNIMES-----SDPTAKAKMAMLI 578
Oy 61 TPAICWIP-FAVAPLSLLQVEIPGTITSVVIGYSAINALPILYLTTRPKE---- 115
579 TDFLCVPISPISFPAISASLKVPLITVSKLILVLFPYPCNSANPLVLAIFKNRRDVI 638
Oy 116 MHRFWNHYRQKSDSKGI-RKMLHSHSGKCHC 151
Db 639 LSKFGCYEVOAQTYRSTSTSTAHNHFPRNGHC 671

RESULT 9
177463
I:uniting hormone/chorionic gonadotropin receptor - rat
C:Species: Rattus sp. (rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: 177463
R:Aalsinki, J.T., Piehla, E.M., Iakkakorpi, J.T., Rajanleht, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A:title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an
A:Reference number: 157668; MUID:92347604; PMID:1353463
A:Accession: 177463
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-700 <RSS>
A:Cross-references: GB:540803; NID:9252163; PIDN:AAB22680.1; PID:9252164
C:Genetics:
A:Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

```

F:153-177/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```
Query Match      16.2%; Score 134; DB 2; Length 700;  
Best Local Similarity 27.2%; Pred. No. 1.3e-05;  
Matches 34; Conservative 31; Mismatches 48; Indels 12; Gaps 4;  
  
QY       1 AQLYSVALFGLINLAIIIVFSYGSMFVSQSATITATEIRNQVKEMILAKREFEIVE 60  
         |||::||:::||::||::||::||::||::||::||::||::||::||  
Db        528 SQQVYLISILL-LNNVAARVVICACIRIYFAQNDELTP-----NKDTKIAKKAILIF 580  
         |||::||::||::||::||::||::||::||::||::||  
QY       61 TDLALCWPI-FVAKPLSLLOVEIPGTTTSWVGIGSALSANLPILYTLLTRPKPE---- 115  
         |||::||::||::||::||::||::||::||::||::||  
Db        581 TDFTCNAPISFFAISAFAKVVELITVTNSKLTLVLFPYNSCANPFLYAIFTKAQFORDELL 640  
         ::||  
QY       116 MIHRF 120  
         ::||  
Db        641 LLSRF 645
```

RESULT 10  
AA9744  
lutropin-choriogonadotropin receptor precursor - rat  
N:Alternate names: lutelizing hormone-choriogonadotropin receptor  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 03-May-1994 #sequence revision 13-Sep-1998 #text\_change 13-Aug-1999  
CAccession: AA9744; AA0545; A41343; A61453; A2460  
R.Koo, Y.B.; Ji, I.; Slaugther, R.G.; Yi, T.H.  
Endocrinology 128, 2297-2308, 1991  
J>Title: Structure of the lutelizing hormone receptor gene and multiple exons of the coc  
A:Reference number: AA9744; MUID:91209270; PMID:2019252  
A:Accession: AA9744  
A>Status: preliminary  
Molecule type: DNA  
A:Molecule type: DNA  
A:Residues: 1-700 <KOQ>  
A:Cross-references: GB:M68917  
A>Note: authors translated the codon CAA for residue 307 as Glu. AAC for residue 355 as I  
R.Tsal-Morris, C.H.; Buczeko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.  
J Biol Chem 266 11355-11359, 1991  
J>Title: Structural organization of the rat lutelizing hormone (LH) receptor gene.  
A:Reference number: AA0545; MUID:91205455; PMID:2040640  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-320 <TSA>  
A:Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923; GB  
R.McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosenblilt, N.; Nikolic, K  
Science 245 494-499, 1989  
J>Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-coupled  
A:Reference number: A41343; MUID:89332512; PMID:2502842  
A:Accession: A41343  
Molecule type: mRNA  
A:Residues: 1-700 <MCP>  
A:Cross-references: GB:M6199; NID:g205178; PID:NAAA1528.1; PID:g205179  
R.Dutau, W.U.; Minegishi, T.; Buczeko, E.S.; Delgado, C.Y.; Zhang, R.  
J Steroid Biochem 33, 715-720, 1989  
J>Title: Characterization and structure of ovarian and testicular LH/hCG receptors.  
A:Reference number: A61453; MUID:90097014; PMID:2601325  
A:Accession: A61453  
A>Status: preliminary  
Molecule type: Protein  
A:Residues: 27-33,'X',35-37,'X',39,'X',41-44 <DUF>  
R.Rocher, P.C.; Ryan, R.J.  
J Biol Chem 264, 4636-4641, 1989  
J>Title: Purification, characterization, and amino-terminal sequence of rat ovarian recep  
A:Reference number: A32460; MUID:89174723; PMID:2925659  
A:Accession: A32460  
Molecule type: Protein  
A:Residues: 27-32,'IX',35-37 <ROC>  
C:Genetics:  
A:Initons: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2  
C:Superfamily: glycoprotein hormone receptor: leucine-rich alpha-2-glycoprotein repeat ho  
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece  
j:54-77/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

```

Query Match      16.2% ; Score 134; DB 2; Length 700;
Best Local Similarity 27.2% ; Pred. No. 1.3e-05;
Matches   34; Conservative    31; Mismatches  48; Indels   12; Gaps     4;

QY          1 AQTYSVAIEFGINLAAPITIVFVSYSGMTSYSHQSATITETIRNOVKKEMILAKRFPIIV 60
              :|:::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db          528 SÖYIITSILI-LNVAVFAVCACAYIRIYAVONPELITP-----NDOTKIAKMAAILP 580

QY          61 TDALCWIPFI-FVAKPLSLQLVELPGTISWVVIGSAINSAINPLIYLTLTRPKF---- 115
              ||::||::||::||::||::||::||::||::||::||::||::||::||:
Db          581 TDFPCMAPISFFAISAARFKVPILTIVTSKLITLVLEFPNASCANPELLAIITKAFQDPFL 640

QY          116 MIHRF 120
              ::||
Db          641 LISRHF 645
```

## A42395

N:lutropin receptor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
 C:Accession: A42395  
 J:Gudermann, T.; Birbaumer, M.; Birbaumer, L.  
 J. Biol. Chem. 267, 4479-4488, 1992  
 A:Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenylyl  
 none receptor expressed for dual coupling of the murine luteinizing hormone receptor to adenylyl  
 A:Reference number: A42395; MUID:92165799; PMID:1311310  
 A:Accession: A42395  
 C:Status: preliminary  
 C:Feature: nucleotide type: mRNA  
 C:Residues: 1-700 <GDU>  
 A:Cross-references: GB:M81310; GB:M81318; NID:g198811; PIDN:AAA3432.1, PID:g198812  
 A:Note: sequence extracted from NCBI backbone (NCBIN:844064, NCBIT:844066)  
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat  
 C:Keywords: G protein-coupled receptor; transmembrane protein  
 F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
 F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
 F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
 F:128-155/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
 F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
 F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
 F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

### Best Local

```

QY      1 AQTYSVAIFLGINLAAFLIIIVSYSGMIFYSHQSALITATEIRNOVKKEMILAKRFVIIV 60
Db      528 SQQYIIISILL-INAAAFVVICACVYRIIFAQNPBELIAP-----NDQTKAKMALIIF 560
QY      61 TDALCIWPI-EVAAKPLSLLOVEIPGTTISWVIGVSAINSALNPILYLTTRPKY--- 115
Db      581 TDTTCWAPISFPAISAFKVPILITVNSVQLLVLEIYYPNNSCANPLTAAVFKFAQDQDFL 640

```

QY	116	MIHRF	120
		::	
Db	641	LSRF	645

## JC7389

Query Match	16.2%;	Score 134;	DB 2;	Length 814;
Best Local Similarity	26.8%;	Pred. No. 1.5e-05;		
Matches 44;	Conservative 32;	Mismatches 60;	Indels 28;	Gaps 6

## Db 624 7

```

QY      1 AQTYSVAFLGINTLMAFLITYVESCSMYSVSHOSALTATEIRNQVKCEKILAKPEFFVF 60
Db      624 AQTYSVLL-LNLTALFLVCACTYKIKYCTVHPNHRSG-----SKDNTAKMAVLIF 678
QY      61 TDAICWLP-LFVAKPLSLIQVEPGLTISWVVGYSAINALNPILYTLTPPK----- 114
Db      677 TDFLCAPISFYMASVADLRPLITYNSKILLVLFYLSNCANPFIYALFTYAFPGDVI 736
QY      115 -----EMTHRFMYNRQRKSRKIRKMLHSSGGCKCHCR 153
Db      737 ILTKVGLCGRRADLR-GGYVSKG-----SSGVCHQGR 770

```

## JC7390

```

Query Match          16.0%; Score 132; DB 2; Length 793;
Best Local Similarity 27.3%; Pred. No. 2.4e-05;
Matches 33; Conservative 28; Mismatches 40; Indels 20; Gaps 4;

OY      1 AATYSVAIFLGIIMLAATITIVFSGSMFYSQMSATATATETINNOVKEMILAKRPFPIVF 60
|||::: : : : |||: : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 598 AQQVYIISVL-LNLAFLVICCTCYKICVAVHNPYCGS-----SKDNIARKMALILF 650

QY 61 TDALCWIP-----FAKPLSLQVIRGTTISWVIGSAINSLNPLIYLTTRPF 113

Db 651 TDFLCMAPIFYMSAVVDRPL-----ITVSNKILLVLPYPLNSCANPFLYALFTKAKF 704

QY 114 K 114

Db 705 R 705

RESULT 14

JC2237

Follicotropin receptor, testis - horse

N:Alternate names: eFSHR

C:Species: Equus caballus (domestic horse)

C:Date: 26-Aug-1985 #sequence, revision 07-Oct-1994 #text\_change 13-Aug-1999

C:Accession: JC2237; JC2370

R:Robert, P.; Amselem, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bidart

Chem. Biophys. Res. Commun. 201, 201-207, 1994

A:Title: Cloning and sequencing of the equine testicular follicotropin receptor.

A:Reference number: JC2237; MUID:94256980; PMID:8198575

A:Accession: JC2237

A:Molecule type: mRNA

A:Residues: 1-694 <ROB>

A:Cross-references: GB:S70150; NID:9546896; PIDN:AAB30854.1; PID:9546897

A:Experimental source: testis

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C:Keyword: glycoprotein; hormone receptor; transmembrane protein

F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:71-96/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:365-386/Domain: transmembrane #status predicted <TM1>

F:398-420/Domain: transmembrane #status predicted <TM2>

F:443-464/Domain: transmembrane #status predicted <TM3>

F:485-507/Domain: transmembrane #status predicted <TM4>

F:528-549/Domain: transmembrane #status predicted <TM5>

F:573-596/Domain: transmembrane #status predicted <TM6>

F:608-629/Domain: transmembrane #status predicted <TM7>

F:191,199,266,293/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.9%; Score 131.5; DB 2; Length 694;

Best Local Similarity 26.5%; Pred. No. 2,3e-05;

Matches 43; Conservative 32; Mismatches 60; Indels 27; Gaps 6;

1 AQTYSVAIFGINLAIFIIVSGSMFYSHQSATITATEIRNQVKEMILAKRFFIIVF 60

Db 526 SOLYVMSL-LVNLVLAFFVYCGCYHIVLRNPNISS-----SSDRIKAKMALILF 578

QY 61 TDALCWIP-----FAKPLSLQVIRGTTISWVIGSAINSLNPLIYLTTRPF 113

Db 579 TDFLCMAPIFYMSAVVDRPL-----ITVSNKILLVLPYPLNSCANPFLYALFTKAKF 704

QY 116 MIHFW-----HNYRQRKSMDSKGRKMLHSSGKGC 151

Db 639 LLSKFCYEMQQLYRTERSSA-----HISPRNGHC 671

RESULT 15

QNHUT

lutropin-choriogonadotropin receptor precursor - human

N:Alternate names: luteinizing hormone-choriogonadotropin receptor

C:Species: Homo sapiens (man)

C:Date: 22-Jan-1993 #sequence, revision 16-Feb-1996 #text\_change 22-Jun-1999

C:Accession: A36243; B36243; A23728; B36120

R:Minagishi, T.; Nakamura, K.; Takakura, Y.; Miyamoto, K.; Hasegawa, Y.; Iwaki, Y.; Igata

Biochem. Biophys. Res. Commun. 172, 1049-1054, 1990

A:Title: Cloning and sequencing of human LH/hCG receptor cDNA.

A:Reference number: A36243; MUID:91058534; PMID:2244890

A:Accession: A36243

A:Molecule type: mRNA

A:Residues: 1-699 <MIN>

A:Cross-references: GB:M3108; NID:9187135; PIDN:AAA59515.1; PID:9187136

A:Experimental source: ovary

A:Accession: B36243

A:Molecule type: mRNA

A:Residues: 1-226,290-699 <M12>

R:Jia, X.C.; Oikawa, M.; Bo, M.; Tanaka, T.; Ny, T.; Boime, I.; Hsueh, A.J.W.

Mol. Endocrinol. 5, 759-768, 1991

A:Title: Expression of human luteinizing hormone (LH) receptor: interaction with LH and

A:Reference number: A23728; MUID:92017881; PMID:1922095

A:Accession: A23728

A:Molecule type: mRNA

A:Residues: 1-123,'R',125-311,'N',313-699 <JIA>

A:Cross-references: GB:S5793

A:Experimental source: ovary granulosa and luteal cells; thyroid

A>Note: 312-Ser was also found; the authors translated the codon UAC for residue 546 as J

R:Prizler, A.L.; Robbins, L.S.; Stork, P.J.; Sprengel, R.; Segaloff, D.L.; Cone, R.D.

Mol. Endocrinol. 4, 1264-1276, 1990

A:Title: Isolation of TSH and LH/CG receptor cDNAs from human thyroid: regulation by tsh

A:Reference number: A36120; MUID:91155962; PMID:2223030

A:Accession: B36120

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-6,'P',8-18,'A',20-26,'R',29-43,'A',45,'A',47,'RPS',52-67,'S',69-123,'R',12

A:Cross-references: GB:M73746; NID:9903745; PIDN:AAA70231.1; PID:9903746

A:Experimental source: thyroid

R:Zhu, H.; Wang, H.; Ascoli, M.

Mol. Endocrinol. 9, 141-150, 1995

A:Title: The lutropin/choriogonadotropin receptor is palmitoylated at intracellular cyst

A:Reference number: A57564; MUID:95295727; PMID:7776964

A:Contents: annotation; mutation analysis of palmitate binding sites

C:Genetics:

A:Gene: GDB:LMHGR

A:Cross-references: GDB:125260; OMIM:152790

A:Map position: 2p21-2p21

A:Introns: 226/3

A>Note: The exact position of the intron cannot be determined from the experimental data

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece

F:1-699/Product: lutropin-choriogonadotropin receptor precursor, long splice form #statu

F:1-226,290-699/Product: lutropin-choriogonadotropin receptor precursor, short splice fo

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-64/Product: lutropin-choriogonadotropin receptor #status predicted <MAT>

F:23-363/Domain: extracellular hormone binding #status predicted <EHB>

F:50-73/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:74-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:124-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:149-173/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:176-197/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:198-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:364-385/Domain: transmembrane #status predicted <TM1>

F:396-418/Domain: transmembrane #status predicted <TM2>

F:441-462/Domain: transmembrane #status predicted <TM3>

F:483-505/Domain: transmembrane #status predicted <TM4>

F:526-547/Domain: transmembrane #status predicted <TM5>

F:571-594/Domain: transmembrane #status predicted <TM6>

F:606-627/Domain: transmembrane #status predicted <TM7>

F:645-699/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:99,174,195,291,299,313/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:561,677/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:643,644/Binding site: palmitate (Cys) (covalent) #status experimental

F:670/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 15.9%; Score 131; DB 1; Length 699;

Best Local Similarity 29.6%; Pred. No. 2.6e-05;

Matches 34; Conservative 25; Mismatches 48; Indels 8; Gaps 3;

1 AQTYSVAIFGINLAIFIIVSGSMFYSHQSATITATEIRNQVKEMILAKRFFIIVF 60

Mon Feb 3 15:52:40 2003

us-10-049-568-2.rpr

Page 7

Db 524 SOVYLITLIL-INVAFETIICACIKIKYEAIRNELMAT-----NDOTIAKAMAILIF 576

Qy 61 TDALCMIRI-EVAPKPLSLQVEIPGIIISWISYASAINSLNELYLTITRPEK 114

Db 577 TDFTCMAISFPFALSAPKVPFLITVNSKVLLVLFYFIPNSCANPFLAITRTKIQ 631

```
Search completed: February 3, 2003, 14:22:39
Job time : 38 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:03:02 ; Search time 21 Seconds

(without alignments)  
310.085 Million cell updates/sec

Title: US-10-049-568-2

Sequence: 1 AQTYSVAIFLGINIAAFIIT.....HMLHSSGKGCRCRCHLS 157

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	638	77.2	757	1 LGR7_HUMAN	Q9hbx9 homo sapien
2	409	49.5	737	1 LGR8_MOUSE	Q91z25 mus musculu
3	392.5	47.5	754	1 LGR8_HUMAN	Q6wx40 homo sapien
4	207.5	25.1	1115	1 GPCR_LYMT	P46023 lymaea sta
5	143.5	17.4	695	1 FSHR_SHEEP	P35379 oviss aries
6	143	17.3	925	1 GHR_ANTPE	P35409 anthopleura
7	141.5	17.1	695	1 FSHR_MACPA	P32212 macaca fasc
8	140.5	17.0	696	1 FSHR_PIG	P16582 sus scrofa
9	139.5	16.9	692	1 FSHR_PAT	P20395 ratuiss norv
10	139.5	16.9	695	1 FSHR_BOVIN	P35376 bos taurus
11	139	16.8	676	1 LSHR_CAJLA	O02721 callithrix
12	136.5	16.5	695	1 FSHR_PIG	P49059 sus scrofa
13	134	16.2	700	1 LSHR_MOUSE	P30730 mus musculu
14	134	16.2	700	1 LSHR_PAT	P16233 ratuiss norv
15	132.5	16.0	692	1 FSHR_MOUSE	P35378 mus musculu
16	132	16.0	701	1 LSHR_BOVIN	Q28005 bos taurus
17	131.5	15.9	694	1 FSHR_HORSE	P47799 equus caball
18	131	15.9	693	1 FSHR_CHICK	P29763 gallus gall
19	131	15.9	699	1 LSHR_HUMAN	P22888 homo sapien
20	129	15.6	687	1 FSHR_EQUUS	Q95179 equus asinu
21	128.5	15.6	619	1 AA3R_MOUSE	O61618 mus musculu
22	128	15.4	320	1 AA3R_PAT	P28647 ratuiss norv
23	127	15.4	764	1 TSHR_SHEEP	P56495 oviss aries
24	125.5	15.2	695	1 FSHR_HUMAN	P21945 homo sapien
25	124	15.0	338	1 P21X_HUMAN	Q15391 homo sapien
26	124	15.0	763	1 TSHR_BOVIN	Q27987 bos taurus
27	124	15.0	764	1 TSHR_PAT	P21463 ratuiss norv
28	123	14.9	340	1 AA2B_CHICK	O13076 gallus gall
29	123	14.9	366	1 LSHR_CHICK	Q90674 gallus gall
30	122	14.8	764	1 TSHR_CANFA	P14763 canis famli
31	121	14.6	764	1 TSHR_MOUSE	P47750 mus musculu
32	119.5	14.5	332	1 AA2B_PAT	P28276 ratuiss norv
33	119.5	14.5	378	1 EDG3_HUMAN	Q99500 homo sapien

34	119.5	14.5	417	1 AA1B_CANFA	P11615 canis famli
35	119	14.4	519	1 AA1B_HUMAN	P35368 homo sapien
36	118.5	14.3	317	1 AA3R_SHEEP	P35342 oviss aries
37	118.5	14.3	338	1 P21X_MOUSE	O60614 mus musculu
38	118	14.3	332	1 AA2B_MOUSE	P60614 mus musculu
39	117.5	14.2	514	1 AA1B_MOUSE	P27717 mus musculu
40	117.5	14.2	515	1 AA1B_MESAU	P18841 mesocricetu
41	117.5	14.2	515	1 AA1B_PAT	P15823 ratuiss norv
42	117.5	14.2	522	1 NF2F_HUMAN	O955x5 homo sapien
43	117	14.2	326	1 AA1R_BOVIN	P28190 bos taurus
44	117	14.2	326	1 AA1R_CANFA	P47745 cavia porce
45	117	14.2	332	1 AA2B_HUMAN	P29275 homo sapien

## ALIGNMENTS

RESULT 1  
LGR7\_HUMAN STANDARD; PRT; 757 AA.  
ID LGR7\_HUMAN  
AC Q9HXB9;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Relaxin receptor 1 (leucine-rich repeat-containing G protein-coupled receptor 7).  
GN LGR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.  
RX MEDLINE=20388592; PubMed=10935549;  
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,  
RA van der Spek P.J., van Duin M., Heuvel A.J.W.;  
RT "The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7";  
RT Mol. Endocrinol. 14:1257-1271(2000).  
RL  
RN [2]  
RP CHARACTERIZATION.  
RX PubMed=11517286;  
RA Bartesch O., Bartlick B., Iwell R.;  
RT "Relaxin signalling links tyrosine phosphorylation to phosphodiesterase and adenylyl cyclase activity";  
RL Mol. Hum. Reprod. 7:799-809(2001).  
CC  
CC -1- FUNCTION: Receptor for relaxin. The activity of this receptor is mediated by G proteins leading to stimulation of adenylyl cyclase and an increase of cAMP. Binding of the ligand may also activate a tyrosine kinase pathway that inhibits the activity of a phosphodiesterase that degrades cAMP.  
CC  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC  
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.  
CC  
CC -1- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis, placenta, uterus, ovary, adrenal, prostate, skin and heart. Not detected in spleen.  
CC  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC  
CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLAS A DOMAIN.  
CC  
CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: A0190500; A017167.1; -.  
CC DR MIM: 60654; -.  
CC DR HSP: P01130; ILDR.

```
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR002172; LDL_Recept_A.
DR InterPro: IPR001611; LRR_Nterm.
DR InterPro: IPR00372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00057; 1dl_recept_a; 1.
DR Pfam: PF00560; LRR; 8.
DR PRINTS: PF00237; GPCR_Rhodopsn.
DR SMART: SM00192; LDLA; 1.
DR SMART: SM00370; LRR; 7.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 10.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE: PS01209; LDLA_1; 1.
DR PROSITE: PS00068; LDLA_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
Leucine-rich repeat; Alternative splicing;
DOMAIN 1 409
TRANSMEM 410 430
EXTRACELLULAR (POTENTIAL).
DOMAIN 431 433
CYTOPLASMIC (POTENTIAL).
TRANSMEM 444 464
EXTRACELLULAR (POTENTIAL).
DOMAIN 465 486
EXTRACELLULAR (POTENTIAL).
TRANSMEM 487 507
CYTOPLASMIC (POTENTIAL).
DOMAIN 508 527
CYTOPLASMIC (POTENTIAL).
TRANSMEM 528 548
EXTRACELLULAR (POTENTIAL).
DOMAIN 549 577
EXTRACELLULAR (POTENTIAL).
TRANSMEM 578 598
CYTOPLASMIC (POTENTIAL).
DOMAIN 599 629
CYTOPLASMIC (POTENTIAL).
TRANSMEM 630 650
EXTRACELLULAR (POTENTIAL).
DOMAIN 651 672
EXTRACELLULAR (POTENTIAL).
TRANSMEM 673 757
CYTOPLASMIC (POTENTIAL).
DOMAIN 757 763
LDL-RECEPTOR CLASS A.
TRANSMEM 125 148
LRR 1.
TRANSMEM 150 172
LRR 2.
TRANSMEM 173 196
LRR 3.
TRANSMEM 198 220
LRR 4.
TRANSMEM 221 244
LRR 5.
TRANSMEM 246 269
LRR 6.
TRANSMEM 270 293
LRR 7.
TRANSMEM 294 317
LRR 8.
TRANSMEM 319 341
LRR 9.
TRANSMEM 343 365
LRR 10.
DISULFID 485 563
BY SIMILARITY.
CARBOHYD 36 36
N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 127 127
N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 264 264
N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 272 272
N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 325 325
N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 368 368
N-LINKED (GLCNAC. . .) (POTENTIAL).
VARSPIC 63 96
GDNNGSMQDFKXFASYKTSQYPEASTPECL -> V
(IN ISOFORM 2).
MUTAGEN 637 637
D->Y; LEADS TO CONSTITUTIVE INCREASE OF
BASAL CAMP.
SQ SEQUENCE 757 AA; 86592 MW; 8079B8DF3A3FE21 CRC64;
```

```
Query Match 77.2%; Score 638; DB 1; Length 757;
Best Local Similarity 94.8%; Pred. No. 7.2e-51;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
DB 696 WNYRORXKMSDKG 709
RESULT 2
ID LG8R MOUSE STANDARD; PRT; 737 AA.
AC 091275.
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Relaxin receptor 2 (leucine-rich repeat-containing G protein-coupled
DE receptor 8) (G protein-coupled receptor affecting testicular descent)
DE (G protein-coupled receptor 106).
GN LG8R OR GREAT OR GPR106.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21250990; PubMed=11353515;
RA Overbeek P.A., Gorlov I.P., Sutherland R.W., Houston J.B.,
RA Harrison W.R., Beutiger-Tong H.L., Bishop C.E., Agounik A.I.;
RT "A transgenic insertion causing cryptorchidism in mice."
RL Genesis 30:26-35(2001).
CC - FUNCTION: Receptor for relaxin. The activity of this receptor is
CC mediated by G proteins leading to stimulation of adenylate cyclase
CC and an increase of cAMP. May also be a receptor for Leydig
CC insulin-like peptide (INSL3) (By similarity).
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: Expressed in embryonic and adult gonads of
CC males and females, as well in male gubernaculum. Expressed also
CC in brain. Not detected in kidney, spleen and heart.
CC - DISEASE: Defects in LG8R seems to be a cause of impaired
CC testicular descent (known as cryptorchidism).
CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC - SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC - SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC - MISCELLANEOUS: By homology with the human sequence, it is
CC uncertain whether Met-1 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: AF346501; AAL08943.1; -.
DR GDB: WGI:2153463; Gp106.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR002172; LDL_Recept_A.
DR InterPro: IPR001611; LRR_Nterm.
DR InterPro: IPR00372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00057; 1dl_recept_a; 1.
DR Pfam: PF00560; LRR; 8.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_1;
DR PROSITE: PS01209; LDLA_1; 1.
DR PROSITE: PS00068; LDLA_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
leucine-rich repeat.
KW DOMAIN 1 399
TRANSMEM 400 420
EXTRACELLULAR (POTENTIAL).
DOMAIN 421 438
CYTOPLASMIC (POTENTIAL).
TRANSMEM 439 459
EXTRACELLULAR (POTENTIAL).
DOMAIN 460 490
EXTRACELLULAR (POTENTIAL).
TRANSMEM 491 511
CYTOPLASMIC (POTENTIAL).
DOMAIN 512 520
CYTOPLASMIC (POTENTIAL).
TRANSMEM 521 541
EXTRACELLULAR (POTENTIAL).
```



FT DOMAIN 542 575 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 576 596 5 (POTENTIAL).  
 FT DOMAIN 597 622 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 623 643 6 (POTENTIAL).  
 FT DOMAIN 644 644 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 645 665 7 (POTENTIAL).  
 FT DOMAIN 666 737 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 27 64 LDL-RECEPTOR CLASS A.  
 FT REPEAT 119 142 LRR 1.  
 FT REPEAT 144 166 LRR 2.  
 FT REPEAT 167 190 LRR 3.  
 FT REPEAT 192 214 LRR 4.  
 FT REPEAT 215 238 LRR 5.  
 FT REPEAT 239 262 LRR 6.  
 FT REPEAT 263 286 LRR 7.  
 FT REPEAT 287 310 LRR 8.  
 FT REPEAT 312 314 LRR 9.  
 FT REPEAT 335 358 LRR 10.  
 FT DISULFID 478 556 BY SIMILARITY.  
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 737 AA; 82943 MW; AF60F635EALACE49 CRC64;

Query Match 49.5%; Score 409; DB 1; Length 737;  
 Best Local Similarity 58.9%; Pred. No. 5.7e-30;  
 Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSVAIFLGINIAFIIVSYGSMFYVQSAITATEIRNVKKEMLIAKFPFIVPTDA 63  
 Db 572 YSLGIPGVNLLAFIVITISYVTFMCSHKTLQTAVRSHIGKEVAVARFFETVSDA 631  
 QY 64 LCMPIPIPAKLSLLQVEIPGTITSVVIGSAINSALNPLITLTTPPKEMIRHWHN 123  
 Db 632 ICMPIPVFKILSLQVEIPGTITSVVIFPLPVNSALNPLITLTTPPKKQLKQLHK 691  
 QY 124 YROR 127  
 Db 692 HRRK 695

RESULT 3  
 LG88 HUMAN STANDARD; PRT; 754 AA.  
 Q8WKO;  
 15-JUN-2002 (Rel. 41, Created)  
 15-JUN-2002 (Rel. 41, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled receptor 8) (G protein-coupled receptor affecting testicular descent).  
 GN LG88 OR GREAT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-647.  
 RX MEDLINE=21669315; PubMed=11809971;  
 RA Hsu S.Y., Nakabayashi K., Nishi S., Kumagai J., Kudo M.,  
 RA Sherwood O.D., Hsueh A.J.W.;  
 RT "Activation of orphan receptors by the hormone relaxin.";  
 RL Science 295:671-674(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gorlov I.P., Kamat A., Jones E., Lamb D., Truong A., Bogatcheva N.,  
 RA Bishop C.B., McElreavey K., Agoulnik A.I.;  
 RT "Mutations of the GREAT gene cause cryptorchidism.";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Receptor for relaxin. The activity of this receptor is mediated by G proteins leading to stimulation of adenylate cyclase

CC and an increase of cAMP. May also be a receptor for Leydig  
 CC Insulin-like peptide (INSI3).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed mainly in the brain kidney, muscle,  
 CC testis, thyroid, uterus, peripheral blood cells and bone marrow.  
 CC -!- DISEASE: Defects in LG88 seems to be a cause of impaired  
 CC testicular descent (known as cryptorchidism). It is one of the  
 CC most frequent congenital abnormalities in humans, involving 2-5%  
 CC of male births. Cryptorchidism is associated with increased risk  
 CC of infertility and testicular cancer.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-18 is the initiator.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; AF403384; AAL69324.2; -;  
 CC EMBL; AF453828; AAL73946.1; -;  
 CC MIM; 606655; -;  
 CC MIM; 219050; -;  
 CC DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 CC DR InterPro: IPR002172; LDL\_Recept\_A.  
 CC DR InterPro: IPR001611; LRR.  
 CC DR InterPro: IPR003591; LRR\_Typ.  
 CC DR Pfam; PF00001; 7tm\_1; 1.  
 CC DR Pfam; PF00057; 1d1\_recept\_a; 1.  
 CC DR Pfam; PF00560; LRR\_8.  
 CC DR PRINTS; PR00237; GPCR\_RHODOPSIN.  
 CC DR PRINTS; PR00019; LEUCINRPT.  
 CC DR SMART; SM00192; LDLA\_1.  
 CC DR SMART; SM00369; LRR\_Typ\_10.  
 CC DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE\_NEG.  
 CC DR PROSITE; PS02662; G-PROTEIN RECEPTOR FL2; 1.  
 CC DR PROSITE; PS01209; LDLRA\_1; 1.  
 CC DR PROSITE; PS00068; LDLRA\_2; 1.  
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;  
 CC Leucine-rich repeat.  
 FT DOMAIN 1 416  
 FT TRANSSEM 417 437 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 438 455 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 456 476 2 (POTENTIAL).  
 FT DOMAIN 477 507 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 508 528 3 (POTENTIAL).  
 FT DOMAIN 529 537 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 538 558 4 (POTENTIAL).  
 FT DOMAIN 559 592 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 593 613 5 (POTENTIAL).  
 FT DOMAIN 614 639 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 640 660 6 (POTENTIAL).  
 FT DOMAIN 661 673 EXTRACELLULAR (POTENTIAL).  
 FT TRANSSEM 674 694 7 (POTENTIAL).  
 FT DOMAIN 695 754 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 754 81 LDL-RECEPTOR CLASS A.  
 FT DOMAIN 81 159 LRR 1.  
 FT REPEAT 136 159 LRR 2.  
 FT REPEAT 161 183 LRR 3.  
 FT REPEAT 184 207 LRR 4.  
 FT REPEAT 209 231 LRR 5.  
 FT REPEAT 233 255 LRR 6.  
 FT REPEAT 256 279 LRR 7.  
 FT REPEAT 281 303 LRR 8.  
 FT REPEAT 304 327 LRR 9.  
 FT REPEAT 329 351 LRR 10.  
 FT DISULFID 352 375 BY SIMILARITY.  
 FT CARBOHYD 495 573 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 138 N-LINKED (GLCNAC. . .) (POTENTIAL).



```

FT DISULFID 322 339 BY SIMILARITY.
FT DISULFID 334 352 BY SIMILARITY.
FT DISULFID 346 361 BY SIMILARITY.
FT DISULFID 367 379 BY SIMILARITY.
FT DISULFID 374 392 BY SIMILARITY.
FT DISULFID 386 401 BY SIMILARITY.
FT DISULFID 406 418 BY SIMILARITY.
FT DISULFID 431 431 BY SIMILARITY.
FT DISULFID 425 440 BY SIMILARITY.
FT DISULFID 446 458 BY SIMILARITY.
FT DISULFID 453 474 BY SIMILARITY.
FT DISULFID 465 483 BY SIMILARITY.
FT DISULFID 488 500 BY SIMILARITY.
FT DISULFID 495 513 BY SIMILARITY.
FT DISULFID 507 523 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1115 AA; 125865 MW; E9BB012978CB356C CRC64;

```

```

Query Match 25.1%; Score 207.5; DB 1; Length 1115;
Best local similarity 33.1%; Pred. No. 2.2e-11;
Matches 51; Conservative 25; Mismatches 45; Indels 33; Gaps 5;

```

```

QY 4 YSAVALFGLNIAFLIIFPSVSGMFYSV---HQSATITAEIRNVCYKEMILAKRFEFIVF 60
DB 938 YSVAFLNLTLSFVLSSSYLMF-SVAKKTRSAVRRAESGN---DNAMARNTLIVM 992
QY 61 TDALCMPIFVAKPPLSLQVEIPGTITSWVVGYSANALPILYTTTRPF----- 113
DB 993 TDFCWCVPITIVGFVSLAGARDDQVMIAVFLPLNSATNPVYITLSTAPFLGNVAKR 1052
QY 114 -----KEMTHRF-----WNYRCKRS 129
DB 1053 ANRFKRSFIHSFTGDTKHSYVDGTTSHSYCEKRS 1086

```

```

RESULT 5
FSHR_SHEEP STANDARD; PRT; 695 AA.
ID P35379; Q28573; Q28574; Q9TS19;
01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin
  receptor).
GN FSHR.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
NCBI_TaxID=9940;
[1]
RP SEQUENCE FROM N.A. (ISOFORM FSH-R1).
RC TISSUE=Testis;
RX MEDLINE=93351750; PubMed=8394255;
RA Varney T.A., Saitam M.R., Khan H., Ravindranath N., Payne S.,
RA Seidah N.G.;
RA "Molecular cloning and expression of the ovine testicular follicle
  stimulating hormone receptor."
RL Mol. Cell. Endocrinol. 93:219-226(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS FSH-R4 AND FSH-R3).
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX MEDLINE=93176195; PubMed=8439338;
RA Khan H., Varney T.A., Saitam M.R.;

```

```

RT "Cloning of alternatively spliced mRNA transcripts coding for variants
  of ovine testicular follictrophin receptor lacking the G protein
  coupling domains.";
RL Biochem. Biophys. Res. Commun. 190:888-894(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX MEDLINE=98031015; PubMed=9364440;
RA Varney T.A., Jiang L., Khan H., Macdonald E.A., Laird D.W.,
RA Saitam M.R.;
RA "Molecular cloning, structure, and expression of a testicular
  follictrophin receptor with selective alteration in the carboxy terminus
  that affects signaling function.";
RL Mol. Reprod. Dev. 48:458-470(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
RC STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
RX MEDLINE=20391225; PubMed=10527886;
RA Babu P.S., Jiang L., Saitam A.M., Touyz R.M., Saitam M.R.;
RT "Structural features and expression of an alternatively spliced growth
  factor type I receptor for follictrophin signaling in the developing
  ovary.";
RL Mol. Cell Biol. Res. Commun. 2:21-27(1999).
RN [5]
RP SEQUENCE OF 1-51 FROM N.A.
RX MEDLINE=98031017; PubMed=9364442;
RA Saitam M.R., Subbarayan V.S.R.;
RT "Characterization of the 5' flanking region and potential control
  elements of the ovine follictrophin receptor gene.";
RL Mol. Reprod. Dev. 48:480-487(1997).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
  of isoform FSH-R1 is mediated by G proteins which activate
  adenylyate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
  this does not result in activation of adenylyate cyclase. Isoform
  FSH-R3 may be involved in calcium signaling.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
  (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2,
  FSH-R3 and FSH-R4; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
  testis, but not in kidney.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation-
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
  CC or send an email to license@isb-sdb.ch).
-----
DR EMBL; L07302; AAA31525.1; -
DR EMBL; L12766; AAA31523.1; -
DR EMBL; L12767; AAA31524.1; -
DR EMBL; L36115; AAK70667.1; -
DR EMBL; AJ0131735; CA10495.1; -
DR EMBL; AF090438; AAC61749.1; -
DR HSP; P23945; JCI493.
DR HSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodopn.
DR InterPro; IPR001611; LRR_Nterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_4.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SMO0013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
  KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT SIGNAL 1 17 POTENTIAL.

```

```

RT LH/CG receptor family from mammals." ;
RL Biochem. Biophys. Res. Commun. 197;1062-1069(1993).
CC -1- FUNCTION: PROBABLE RECEPTOR FOR A GLYCOPROTEIN HORMONE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/FSH SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z28332; CAA82186.1; -.
CC PIR; S41908; S41908.
CC PIR; JC2033; JC2033.
CC InterPro; IPR000276; GPCR_Rhodpam.
CC InterPro; IPR001832; GPCR_secretin.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003592; LRR_out.
CC Pfam; PF00001; 7tm_1; 1.
CC Pfam; PF00560; LRR; 6.
CC SMART; SMO0370; LRR; 2.
CC PROSITE; PS00237; G-PROTEIN RECP F1 1; 1.
CC PROSITE; PS50262; G-PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Alternative splicing; Repeat.
CC SIGNAL
CC 1
CC 27
CC 28 925
CC CHAIN
CC -----
CC DOMAIN 28 529
CC TRANSSEM 530 551
CC DOMAIN 552 561
CC TRANSSEM 562 584
CC DOMAIN 585 606
CC TRANSSEM 607 628
CC DOMAIN 629 651
CC TRANSSEM 652 673
CC DOMAIN 674 691
CC TRANSSEM 692 712
CC DOMAIN 713 739
CC TRANSSEM 740 763
CC DOMAIN 764 774
CC TRANSSEM 775 795
CC DOMAIN 796 925
CC DOMAIN 333 461
CC REPEAT 333 349
CC REPEAT 350 384
CC REPEAT 385 419
CC REPEAT 420 453
CC REPEAT 454 461
CC CARBOHYD 61 61
CC CARBOHYD 152 152
CC CARBOHYD 212 212
CC CARBOHYD 435 435
CC CARBOHYD 495 495
CC VASPLIC 235 925
CC MISSING (IN SHORT ISOFORM).
CC SEQUENCE 925 AA; 100059 MW; D03A25636452FBD CRC64;
CC -----
Query Match 17.3%; Score 143; DB 1; Length 925;
Best Local Similarity 33.3%; Prid. No. 1,4e-05;
Matches 37; Conservative 21; Mismatches 47; Indels 6; Gaps 3;
CC -----
CC 6 VALILGNIAPFIVTSYSGMFYSVHQSAITPTREINQYKK-EMIAKRFPIVFDAL 64
CC |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
CC 694 VAFILVNGASFISVMTLYGRMLYV---VSGGDMEGAPKRNDKVKAKMALIVFTDML 749
CC |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
CC 65 CWIDI-FVANPLSLQVEIDPGTITSVWVIGSYAINSALNIDLYLTTRPK 114
CC |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
CC 750 CWABIAIFAGLLAAGOTLLTWTQSKILLVFFPPINSICNFFLYAFTKAPK 800
CC |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

ID	FSHR_MACFA	STANDARD;	PRT;	695 AA.
AC	P32212;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).			
GN	FSHR.			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopithecoidea; Macaca.			
OX	NCBI_TaxID=9541;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Testis;			
RF	MEDLINE=94071854; PubMed=7504463;			
RT	Gromoll J., Danbar B., Sharma R.S., Nieschlag E.;			
RT	"Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis."			
RL	Biochem. Biophys. Res. Commun. 196:1066-1072(1993).			
CC	-1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE CYCLASE.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	FSH/LSH/TSH SUBFAMILY.			
CC	-1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a> or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).			
CC	-----			
DR	EMBL; X74454; CAA52463.1; -.			
DR	PIR; S36452; S36452.			
DR	PIR; JN0898; JN0898.			
DR	HSSB; P23945.1XUN.			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR00372; LRR_Nterm.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	Pfam; PF00560; LRR; 3.			
DR	Pfam; PF01462; LRRNT; 1.			
DR	SMART; SMO0013; LRRNT; 1.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.			
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.			
KW	G-protein coupled receptor; Transmembrane; signal;			
KW	Phosphorylation; Repeat; Leucine-rich repeat.			
FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	695	FOLLICLE STIMULATING HORMONE RECEPTOR.
FT	DOMAIN	18	366	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	367	387	1 (POTENTIAL).
FT	DOMAIN	388	398	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	399	421	2 (POTENTIAL).
FT	DOMAIN	422	443	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	444	465	3 (POTENTIAL).
FT	DOMAIN	466	485	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	486	508	4 (POTENTIAL).
FT	DOMAIN	509	528	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	529	550	5 (POTENTIAL).
FT	DOMAIN	551	573	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	574	597	6 (POTENTIAL).
FT	DOMAIN	598	608	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	609	630	7 (POTENTIAL).
FT	DOMAIN	631	695	CYTOPLASMIC (POTENTIAL).

CC	REPEAT	44	68	LRR 1.	
FT	REPEAT	69	93	LRR 2.	
FT	REPEAT	119	143	LRR 3.	
FT	REPEAT	170	192	LRR 4.	
FT	REPEAT	193	216	LRR 5.	
FT	REPEAT	218	240	LRR 6.	
FT	DISULFID	442	517	BY SIMILARITY.	
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	318	318	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SO	SEQUENCE	695 AA;	78343 MW;	0D6A2A33729B5250 CRC64;	
	Query Match	17.1%;	Score 141.5;	DB 1;	Length 695;
	Best Local Similarity	26.5%;	Pred. No. 1.4e-05;		
	Matches	44;	Conservative 32;	Mismatches 63;	Indels 27; Gaps 6;
Oy	1	AQISVAIFGLINLAFLIIVFSYGSMFYSYHOSAIATATEIRNOVKKEMILAKRFPIVF	60		
Db	527	SQLYVMSL-LVLNLAFLAVVICGCTHTHLYVRNNIVSS-----SSDPTIAKMMAMLIIF	579		
Oy	61	TDALCWIPD-FVAKPESLLOVEIPGTTTSVNVIGYSAINSGLNFIYLTITPRKE---	115		
Db	580	TDFICMAPISEFFAISAKIVPLITVSRAKILLVLVFPYDINSCANPFLYAIFTKRFRDFFI	639		
Oy	116	MIHFW-----NHYRQKSMDSKIGIKRKMILHSSGGKCGCRCH	155		
Db	640	LISKFGCYEMQAIYRTETSTSA-----HNSHPNGHCSAH	676		
RESULT 8					
LSHR_PIG	ID	LSHR_PIG	STANDARD;	PRT;	636 AA.
AC	P16582;				
DT	01-AUG-1990	(Rel. 15, Created)			
DT	01-AUG-1990	(Rel. 15, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)				
GN	(LSH-R) (lateinizing hormone receptor).				
GN	LHCGR.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID=9623;				
RA	[1]				
RA	SEQUENCE FROM N.A.				
RX	MEDLINE=89332517; PubMed=2502844;				
RA	Medjelle H., Mistrabi M., Atger M., Salese R., Thi M.T.V.H.-L.,				
RA	Jolivet A., Guichon-Mantel A., Sar S., Jallat B., Garnier J.,				
RA	Millotom E.;				
RT	"Cloning and sequencing of porcine LH-hCG receptor cDNA: variants				
RT	lacking transmembrane domain.";				
RL	Science 245:525-528(1998).				
CC	- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGNADOTROPIC HORMONE.				
CC	THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH				
CC	ACTIVATE ADENYLYATE CYCLASE.				
CC	- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B, C AND D; ARE				
CC	PRODUCED BY ALTERNATIVE SPLICING.				
CC	- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	FSH/LSH/ISH SUBFAMILY.				
CC	- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL; M29525; AAA31062.1; -.				
DR	EMBL; M29526; AAA31063.1; -.				



CC EMBL: L22319; AAC37324.1; -.  
 DR HSSP: P23945; 1XUN.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR004511; LRR\_Neerm.  
 DR InterPro: IPR000372; LRR\_Neerm.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR Pfam: PF00560; LRR\_4.  
 DR Pfam: PF01462; LRRT; 1.  
 DR SMART: SM00013; LRRT; 1.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECBE\_F1\_1; 1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECBE\_F2\_1; 1.  
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Phosphorylation; Repeat; Leucine-rich repeat.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 1 695 FOLLICLE STIMULATING HORMONE RECEPTOR.  
 FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 367 387 1 (POTENTIAL).  
 FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 399 421 2 (POTENTIAL).  
 FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 444 465 3 (POTENTIAL).  
 FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 486 508 4 (POTENTIAL).  
 FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 529 550 5 (POTENTIAL).  
 FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 574 597 6 (POTENTIAL).  
 FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 609 630 7 (POTENTIAL).  
 FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 44 68 LRR 1.  
 FT REPEAT 69 93 LRR 2.  
 FT REPEAT 119 143 LRR 3.  
 FT REPEAT 170 192 LRR 4.  
 FT REPEAT 193 216 LRR 5.  
 FT REPEAT 218 240 LRR 6.  
 FT DISULFID 442 517 BY SIMILARITY.  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 695 AA; 78064 MW; 18F9DFEPC046380D CRC64;  
 Query Match 16.9%; Score 139.5; DB 1; Length 695;  
 Best Local Similarity 26.5%; Pred. No. 2, 2e-05;  
 Matches 43; Conservative 33; Mismatches 59; Indels 27; Gaps 6  
 QY 1 AQTYSVALFGLINLAFLITVVSQSGSYSHQSALTATIRINQKKEMILAKKFFPIV 60  
 DB 527 SOLYVMSI-LVLNLAFLVVICGCTYHYLVIRENENTIS-----SSDTKIAKKMAMLI 579  
 QY 61 TPLALCWIDI-FVAKFASLLOVEIPGTLISWVIVGYSAINSAINPLIYLLTTRPKK---- 115  
 DB 580 TDFLCMAIISFPAISASLAKVPLITVSKSKILLVLFVYINSCANPFLYIAFTKNPRDFFI 639  
 QY 116 MIHREW-----HNYRQKSMDSKIRKGMHSHSGSGKCHC 151  
 DB 640 LLSKFGCYEVDQAQTIRSETISSTH-----HNFHFRNHC 672  
 RESULT 11  
 LSHR CALJA  
 ID LSHR CALJA STANDARD; PRT; 676 AA.  
 AC 002721;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)  
 DE (LSHR) (luteinizing hormone receptor).  
 GN LHGR.  
 OS Callitrix jacchus (Common marmoset).  
 CC Eukaryota; Euteleostomi; Chordata; Cnidaria; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Cnidaria; Vertebrata; Euteleostomi;  
 CC Mammalia; Euteleostomi; Chordata; Cnidaria; Vertebrata; Euteleostomi;

CC Callitrix.  
 OX NCBI\_Taxid=9483;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=97307697; PubMed=9165039;  
 RA Zhang F.-P., Rannikko A.S., Manna P.R., Fraser H.M., Huhtaniemi I.T.;  
 RT "Cloning and functional expression of the luteinizing hormone  
 receptor complementary deoxyribonucleic acid from the marmoset monkey  
 leucis: absence of sequences encoding exon 10 in other species."  
 RL Endocrinology 138:2481-2490(1997).  
 CC -1- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGNADOTROPIC HORMONE.  
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH  
 CC ACTIVATE ADENYLATE CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC FSH/LSH/TSH SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U80673; AAB53698.1; -.  
 CC DR HSSP: P22888; 1LUT.  
 CC DR InterPro: IPR000276; GPCR\_Rhodopsin.  
 CC DR InterPro: IPR001611; LRR\_Nterm.  
 CC DR InterPro: IPR000372; LRR\_Nterm.  
 CC DR Pfam: PF00001; 7tm\_1; 1.  
 CC DR Pfam: PF00560; LRR\_2.  
 CC DR SMART: SM00013; LRRT1; 1.  
 CC DR PROSITE: PS00237; G-PROTEIN RECP\_F1\_1; 1.  
 CC DR PROSITE: PS00262; G-PROTEIN RECP\_F1\_2; 1.  
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 CC Phosphorylation; Repeat; Leucine-rich repeat.  
 CC FT SIGNAL 1 29  
 CC FT CHAIN 30 676  
 CC FT  
 CC FT DOMAIN 30 340  
 CC FT TRANSMEM 341 362  
 CC FT TRANSMEM 363 372  
 CC FT TRANSMEM 373 393  
 CC FT TRANSMEM 394 416  
 CC FT TRANSMEM 417 439  
 CC FT TRANSMEM 440 459  
 CC FT TRANSMEM 460 482  
 CC FT TRANSMEM 483 502  
 CC FT TRANSMEM 503 526  
 CC FT TRANSMEM 527 547  
 CC FT TRANSMEM 548 571  
 CC FT DOMAIN 572 582  
 CC FT TRANSMEM 583 604  
 CC FT TRANSMEM 605 676  
 CC FT DOMAIN 677 751  
 CC FT REPEAT 126 151  
 CC FT REPEAT 153 175  
 CC FT REPEAT 176 200  
 CC FT REPEAT 225 248  
 CC FT DISULFID 416 491  
 CC FT CARBOHYD 103 103  
 CC FT CARBOHYD 178 178  
 CC FT CARBOHYD 199 199  
 CC SEQUENCE 676 AA; 75677 MW; FBO1822764154E1 CRC64;  
 CC  
 CC Query Match 16.8%; Score 139; DB 1; Length 676;  
 CC Best Local Similarity 31.3%; Pred. No. 2,4e-05;  
 CC Matches 36; Conservative 24; Mismatches 47; Indels 8; Gaps 3;  
 CC  
 CC 1 AQTYSVALFLGILNLAFLIIIVSGMFWYSVHQSALITAEIRNQVKEMILAKRFPIIV 60

DB 501 SQYITLTLI-LANVAPLITACACIYIPVAPNPELMAI-----NKOTAKKMAILIF 553  
 QY 61 TDALCWIPD-FVAKPSLLQVEIPGTTSMVVGYSAINSAINDILYTTTRPDK 114  
 DB 554 TDFCTMAPISFFAISAFAKMPILITVNSKVLILVLFYPINSCANPELYAIFKTFR 608  
 RESULT 12  
 ID FSHR\_PIG STANDARD; PRT: 695 AA.  
 AC P49059; O77514;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Folicle stimulating hormone receptor precursor (FSH-R) (Folliotropin  
 DE receptor).  
 GN FSHR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Suidae.  
 OX NCBI\_Taxid=9623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=96011644; PubMed=7590277;  
 RA Remy J.J., Lahbib-Mansale Y., Yarle M., Bozon V., Couture L.,  
 RA Pajot E., Grebert D., Salese R.;  
 RT "The porcine follicle-stimulating hormone receptor: cDNA cloning, functional  
 RL expression and chromosomal localization of the gene."  
 RL gene 163:257-261(1995).  
 CC  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Ovary;  
 CC RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,  
 CC RA la Barbera A.R.;  
 CC RL "Porcine follicle-stimulating hormone receptor."  
 CC RL Submitted (SBP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY  
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE  
 CC ADENYLATE CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC FSH/LSH/TSH SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U31966; AAB6933.1; -.  
 CC DR EMBL: AF025377; AAC24981.1; -.  
 CC DR HSSP: P23945; 1XUN.  
 CC DR InterPro: IPR000276; GPCR\_Rhodopsin.  
 CC DR InterPro: IPR001611; LRR\_Nterm.  
 CC DR InterPro: IPR000372; LRR\_Nterm.  
 CC DR Pfam: PF00001; 7tm\_1; 1.  
 CC DR Pfam: PF00560; LRR\_2.  
 CC DR Pfam: PF01463; LRRT1; 1.  
 CC DR SMART: SM00013; LRRT1; 1.  
 CC DR PROSITE: PS00237; G-PROTEIN RECP\_F1\_1; 1.  
 CC DR PROSITE: PS00262; G-PROTEIN RECP\_F1\_2; 1.  
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 CC Phosphorylation; Repeat; Leucine-rich repeat.  
 CC FT SIGNAL 1 17  
 CC FT CHAIN 18 695  
 CC FT DOMAIN 18 366  
 CC FT TRANSMEM 367 387  
 CC FT DOMAIN 388 398  
 CC CYTOPLASMIC (POTENTIAL).









RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE OF 1-51 FROM N.A.  
 RX MEDLINE=93093308; PubMed=1459341;  
 RA Huhantemi I.T., Eskola V., Pakarinen P., Matikainen T.,  
 RA Sprengel R.;  
 RT "The murine luteinizing hormone and follicle-stimulating hormone  
 RT receptor genes: transcription initiation sites, putative promoter  
 RT sequences and promoter activity.";  
 RT Mol. Cell. Endocrinol. 88:55-66(1992).  
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY  
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE  
 CC ADENYLATE CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC FSH/LSH/TSH SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF095642; AAC67559.1; -;  
 CC EMBL; AK016635; BAB30351.1; -;  
 CC EMBL; S49632; AAB24401.1; -;  
 CC EMBL; M87570; AAA37641.1; -;  
 CC MGI; 95583; Behr.  
 CC InterPro; IPR00276; GPCR Rhodopsin.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR000372; LRR\_Nterm.  
 CC Pfam; PF00001; 7tm.1; 1.  
 CC Pfam; PF00560; LRR\_3.  
 CC Pfam; PF01462; LRRNT; 1.  
 CC PRINTS; PR00237; GPCR RHODOPSIN.  
 CC SMART; SM00013; LRRNT; 1.  
 CC PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1;  
 CC PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 CC KMW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 CC phosphorylation; Repeat; Leucine-rich repeat.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 692 FOLLICLE STIMULATING HORMONE RECEPTOR.  
 CC DOMAIN 18 365 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 366 386 1 (POTENTIAL).  
 CC DOMAIN 387 397 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 398 420 2 (POTENTIAL).  
 CC DOMAIN 421 442 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 443 464 3 (POTENTIAL).  
 CC DOMAIN 465 507 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 485 507 4 (POTENTIAL).  
 CC DOMAIN 508 527 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 528 549 5 (POTENTIAL).  
 CC DOMAIN 550 572 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 573 596 6 (POTENTIAL).  
 CC DOMAIN 597 607 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 608 629 7 (POTENTIAL).  
 CC DOMAIN 630 692 CYTOPLASMIC (POTENTIAL).  
 CC REPEAT 44 68 LRR 1.  
 CC REPEAT 69 93 LRR 2.  
 CC REPEAT 119 143 LRR 3.  
 CC REPEAT 170 192 LRR 4.  
 CC REPEAT 193 216 LRR 5.  
 CC REPEAT 218 240 LRR 6.  
 CC DISULFID 441 516 BY SIMILARITY.  
 CC CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 436 436 Q -> K (IN REF. 2).  
 CC SEQUENCE 692 AA; 77769 MW; 4B57229180563A44 CRC64;

Query Match 16.0%; Score 132.5; DB 1; Length 692;  
 Best Local Similarity 26.6%; Pred. No. 9.5e-05;  
 Matches 42; Conservative 30; Mismatches 57; Indels 29; Gaps 5;

QY 1 AQIYVAIFGIMIAAFIIIVFSYSGMEFYSVHOSAITAEIRNOVKEMILAKRFFIIF 60  
 DB 526 SQLYVMAL-LVNLALAFVVICGCTHIYLVTRNPNTVSS-----SRDTKIAKRMATLIF 578  
 QY 61 TDALCWIPF-FVAKPUSLQVEIPGIIITSWVVIQYSAINSALNPILYTLTPPEKE---- 115  
 DB 579 TDFLCWAPILFFAISASLKVPLITVSRAKILLVLFYPINSCANPFLYALFTKNFRDFEV 638  
 QY 116 MIHRF-----WHNYRKSKMDSKGR 136  
 DB 639 LMSKFGCYEQAOIKYKETSITHNFRSRKNPCSSAPR 676

Search completed: February 3, 2003, 14:21:48  
 Job time : 22 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:20:08 ; Search time 73 Seconds  
(without alignments)  
443.142 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 826  
Sequence: 1 AQLYSVAIFLGINLAFIIL.....HMLHSSGGKCHRCRCHLS 157

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp Vertebrate.\*  
14: sp Unclassified.\*  
15: sp\_rvlnus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224.5	27.2	334	5 Q9VYG0	Q9VYG0 drosophila
2	212.5	25.7	359	5 Q9VBP0	Q9VBP0 drosophila
3	154	18.6	693	13 Q9DGC5	Q9DGC5 oreochromis
4	148	17.9	1012	5 Q9SYV6	Q9SYV6 asterina pe
5	148	17.9	1280	5 Q9SYV7	Q9SYV7 asterina pe
6	143.5	17.4	696	13 Q9BTR5	Q9BTR5 ictalurus p
7	139.5	16.9	688	11 Q64183	Q64183 rictus sp.
8	138	16.7	307	13 Q90XC7	Q90XC7 salmo salar
9	134	16.2	814	13 Q91949	Q91949 oncorhynch
10	133.5	16.0	424	13 Q9D63	Q9D63 rana catesb
11	132	16.0	701	6 Q8SP8	Q8SP8 bos taurus
12	132	16.0	793	13 Q91948	Q91948 oncorhynch
13	131	15.9	699	4 Q15996	Q15996 homo sapien
14	131	15.9	701	4 Q14751	Q14751 homo sapien
15	131	15.9	779	13 Q918N7	Q918N7 morone saxa
16	130.5	15.8	778	13 Q98TF4	Q98TF4 oreochromis

17	130	15.7	1050	5 Q9BN18	Q9BN18 drosophila
18	130	15.7	1300	5 Q9NKD6	Q9NKD6 drosophila
19	130	15.7	1360	5 Q9ND11	Q9ND11 drosophila
20	128	15.5	658	13 Q9PVN9	Q9PVN9 oncorhynch
21	127	15.4	662	13 Q9BTR4	Q9BTR4 ictalurus p
22	125.5	15.2	410	4 Q16225	Q16225 homo sapien
23	125	15.1	662	13 Q9PWL6	Q9PWL6 claris gar
24	124	15.0	696	13 Q9DGF5	Q9DGF5 cynops pyr
25	123.5	15.0	695	11 Q8R428	Q8R428 cavia porce
26	123	14.9	335	6 Q6387	Q6387 mustela vis
27	122	14.8	293	6 Q9SMF7	Q9SMF7 canis fami
28	122	14.8	763	6 Q9BGN4	Q9BGN4 felis silve
29	121	14.6	601	13 Q42500	Q42500 melagris g
30	121	14.6	764	13 Q9D697	Q9D697 mus muscul
31	119.5	14.5	384	13 Q9PU08	Q9PU08 fuqu rubrip
32	119.5	14.5	398	4 Q9H228	Q9H228 homo sapien
33	119.5	14.5	515	11 Q9DBL0	Q9DBL0 mus muscul
34	119.5	14.5	724	13 Q9PVP0	Q9PVP0 oncorhynch
35	119	14.4	739	6 Q9BGS5	Q9BGS5 sus scrofa
36	119	14.4	764	6 Q9BGS6	Q9BGS6 sus scrofa
37	119	14.4	764	6 Q8SP9	Q8SP9 sus scrofa
38	117.5	14.2	418	13 Q90K12	Q90K12 xenopus lae
39	117.5	14.2	470	4 Q43200	Q43200 homo sapien
40	117.5	14.2	533	4 Q9P1V4	Q9P1V4 homo sapien
41	116	14.0	885	5 Q61232	Q61232 lymanaea sta
42	115.5	14.0	408	4 Q96RV1	Q96RV1 homo sapien
43	115	13.9	326	4 Q8TAM8	Q8TAM8 homo sapien
44	114.5	13.9	829	5 Q9VEG4	Q9VEG4 drosophila
45	114.5	13.9	831	5 Q94979	Q94979 drosophila

## ALIGNMENTS

Q9VYG0	PRELIMINARY;	PRT;	334 AA.
AC Q9VYG0;			
DT 01-MAY-2000 (TREMREL. 13, Created)			
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)			
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)			
DE CG4187 protein.			
GN CG4187.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophila.			
OX NCBI_TaxID=7227;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=BERKELEY;			
RX MEDLINE=20196006; PubMed=10731132;			
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,			
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA Abail J.F., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,			
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Basley E.M.,			
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA Fustin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,			
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,			
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Nobarty C., Morris J., Moshell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AF003491; AAF48237.1; -.  
 DR FlyBase; FBgn0030458; CG4187.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR InterPro; IPR001901; SecE.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHDOPSN.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_FL\_2; 1.  
 DR PROSITE; PS01067; SECE\_SEC6LG; UNKNOWN\_1.  
 SQ SEQUENCE 334 AA; 37020 MW; 1662F2D467534061 CRC64;  
  
 Query Match 27.2%; Score 224.5; DB 5; Length 334;  
 Best Local Similarity 32.3%; Pred. No. 8.8e-15;  
 Matches 54; Conservative 27; Mismatches 67; Indels 19; Gaps 4;  
  
 QY 4 YSAVAFGLGINTLAFLIIVSYGSMFYVHQS--ALTAETIRNOVKEMILAKRFFIVFT 61  
 Db 166 YSAFLFVNLVNLSTLFLIFSYIRMLQALRDSGGGKRST---HSGREVVATRAIIVTT 221  
  
 QY 62 DALCWIPIFAKPLSLQVEIPGTTISWVIGSAINSLNPILYTLTRPFKMRHF 121  
 Db 222 DCACMLPIIVKIALASCEISPDLYAMLVLPVNSALNPVLTITLTAFFKQLRVC 281  
  
 QY 122 HNY-----RQRKMSDKIRKIMLHSS--GKGCRCRCH 155  
 Db 282 HTLPSCLVNNETRSQTOTAYESGLSVLAHLGGVGGSGRKMSH 328  
  
 RESULT 2  
 Q9VBP0 PRELIMINARY; PRT; 359 AA.  
 Q9VBP0;  
 01-MAY-2000 (TReMBLrel. 13, Created)  
 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE CG5042 protein.  
 GN CG5042 OR CG5046.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Bokoyeva D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibbegan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarty C., Morris J., Moshell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; AB003753; AAF56490.1; -.  
 DR FlyBase; FBgn0039354; CG5042.  
 DR InterPro; IPR002106; AATRNA\_1gaeeit.  
 DR InterPro; IPR00276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHDOPSN.  
 DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_FL\_1; 1.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_FL\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.  
 SQ SEQUENCE 359 AA; 40880 MW; 7191BBED7C9E739 CRC64;  
  
 Query Match 25.7%; Score 212.5; DB 5; Length 359;  
 Best Local Similarity 35.9%; Pred. No. 1.6e-13;  
 Matches 47; Conservative 24; Mismatches 55; Indels 5; Gaps 3;  
  
 QY 3 YSAVAFGLGINTLAFLIIVSYGSMFYVHQSATATETIRNOVKEMILAKRFFIVFTD 62  
 Db 199 YSAFAVFLGVNLLVWIMLMYALISITWTR--SATPL--TLDDCEFAVAFPIVLTLD 254  
  
 QY 63 ALCWIPIFAKPLSLQVEIPGTTISWVIGSAINSLNPILYTLTRPFKMRHF 121  
 Db 255 FLCWPIIWKIWFVFNYSIDIIYAMLVFVLPVNSALNPVLTITLTAFFKQLRVC 314  
  
 QY 122 HNYRQRKMSDKIRKIMLHSS--GKGCRCRCH 152  
 Db 315 KITSKRAEA 325  
  
 RESULT 3  
 Q9DGC5 PRELIMINARY; PRT; 693 AA.  
 Q9DGC5;  
 01-MAR-2001 (TReMBLrel. 16, Created)  
 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE G-protein coupled receptor II.  
 GN TGRH-R11.  
 OS *Oreochromis niloticus* (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;  
 OC Cichlidae; Oreochromis.  
 OC NCB1\_TaxID=8128;

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Oba Y., Hirai T., Yoshinura Y., Yao Z., Nagahama Y.;
RT      "Hillapla gonadotropin receptor II."
RL      Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB041763; BAB16107.1; -.
DR      HSSP; P22688; ILUT.
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      InterPro; IPR001611; LRR.
DR      Pfam; PF00001; 7tm_1; 1.
DR      Pfam; PF00560; LRR; 2.
DR      PRINTS; PR00237; GPCR_RHODOPS.
DR      PROSITE; PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.
DR      PROSITE; PS50262; G_PROTEIN_RECPE_F1_2; 1.
KW      Receptor.
SQ      SEQUENCE      693 AA;  77856 MW;  C3BBEFD8BCFC8988 CRC64;

Query Match      18.6%; Score 154; DB 13; Length 693;
Best Local Similarity 34.5%; Pred. No. 2.8e-07;
Matches 40; Conservative 24; Mismatches 44; Indels 8; Gaps 4;

QY      1 AQLVSALFGLINLAFLIIIVFSYGSMEYSVHQSATITAEIRNOYKEMILAKREFPIV 60
DB      524 AQLVVAVLILINVAFLIVCCYICITLSVNHPESTR--RGDKR----IAKRAVAVLIF 576

QY      61 TDAICWIFI-EVAKSLSLQVEIPGTITSWVIGYSAINSLNPLIYLTITRPPK 115
DB      577 TDFLCMAEISFPAISALALRMLPITVSHSKILILFYPINSLCNPLIYITFAFRK 632

RESULT 4
Q95Y16 PRELIMINARY; PRT; 1012 AA.
Q95Y16
AC      Q95Y16;
ID      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Glycoprotein hormone receptor.
GN      APGHR.
OS      Ascarina pectinifera (Starfish).
OC      Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC      Asteroidea; Valvatacea; Valvataida; Asterinidae; Asterina.
OX      NCBI_TaxId=7594;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT      "cDNA cloning and functional analysis of a novel member of the
RT      glycoprotein hormone receptor family from a starfish Asterina
RT      pectinifera."
RL      Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB061862; BAB68209.1; -.
DR      InterPro; IPR002106; ALCRNA_1class1.
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      InterPro; IPR001611; LRR.
DR      InterPro; IPR000372; LRR_Nterm.
DR      Pfam; PF00001; 7tm_1; 1.
DR      Pfam; PF00560; LRR; 14.
DR      SMART; SM00013; LRRNT; 1.
DR      PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR      PROSITE; PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.
DR      PROSITE; PS50262; G_PROTEIN_RECPE_F1_2; 1.
KW      Receptor.
SQ      SEQUENCE      1012 AA;  112623 MW;  52A70E7A88C46E0A CRC64;

Query Match      17.9%; Score 148; DB 5; Length 1012;
Best Local Similarity 32.9%; Pred. No. 1.7e-06;
Matches 51; Conservative 27; Mismatches 51; Indels 26; Gaps 9

QY      1 AQLVSALFGLINLAFLIIIVFSYGSMEYSVHQS-ALTATEIRNOYKEMILAKREFPIV 59
DB      719 AKYV-VESIIILINLAFLIIVIMACYASITYLAIQSHANWCNDSR-----VARRMSLIV 769

QY      60 FTDAICWIFI-EVAKSLSLQVEIPGTITSWVIGYSAINSLNPLIYLTITRPPK 114

```

Dd	770	FTDPAACVAFPIFFSLTAAGRLISLDGAKVLTFVL-----PLNSCANPFLYTLITLKQFK	825
Qy	115	---EMHRRFWHN--YRORKSMDSKGRKMLHHSS	144
Dd	826	KDKCTIMRSLSNRVFRQR-SM-SRSITLILSGRHPS	858
 RESULT 5			
ID	Q95YT7	PRELIMINARY; PRT; 1280 AA.	
AC	Q95YT7		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Glycoprotein hormone receptor.		
GN	APOBHR.		
OS	Asterina pectinifera (Starfish).		
OC	Euteleostomi; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;		
CC	Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.		
NCBI_TaxID=7594;	[1]		
RN			
RP	SEQUENCE FROM N.A.		
RA	Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.:		
RT	"cDNA cloning and functional analysis of a novel member of the		
RT	glycoprotein hormone receptor family from a starfish Asterina		
RT	pectinifera."		
RU	Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL, AB061861; BAB68208.1; -		
DR	InterPro; IPR002106; AAKRNA_1ligaseII.		
DR	InterPro; IPR000276; GPCR_Rhodopsn.		
DR	InterPro; IPR001611; LRR.		
DR	InterPro; IPR000372; LRR_Nterm.		
DR	Pfam; PF00560; LRR_14.		
DR	SMART; SMO0013; LRRNT; 1.		
DR	PROSITE; PS00339; AA_TRNA_LIGASE_I_2; UNKNOWN_1.		
DR	PROSITE; PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.		
DR	PROSITE; PS00262; G_PROTEIN_RECPE_F1_2; 1.		
KW	Receptor.		
SQ	SEQUENCE 1280 AA; 141700 MW; 0AB0ECCCDD80BA CRC64;		
 Query Match 17.9%; Score 148; DB 5; Length 1280; Best Local Similarity 32.9%; Pred.No.2.2e-06; Matches 51; Conservative 27; Mismatches 51; Indels 26; Gaps 9;			
Qy	1	AQIYSVAIFLGINLAFFIVFSYGSMFVSQGS-AITATETIRNQVKEMILAKRPFPIV	59
Dd	719	AKRY-VGSILINIAFIAMTICASYIAYAIQSHAHMCNDSR-----VARRMSLIIV	769
Qy	60	FPTALMWIP-----FVAPDLILOVEIRPGTITSWVIGSYAINSALNPILYTLTPRPK	114
Dd	770	FDPACMAPIAFFSLTAAGRLISLDGAKVLTFVL-----PLNSCANPFLYTLITLKQFK	825
Qy	115	---EMHRRFWHN--YRORKSMDSKGRKMLHHSS	144
Dd	826	KDKCTIMRSLSNRVFRQR-SM-SRSITLILSGRHPS	858
 RESULT 6			
ID	Q98T85	PRELIMINARY; PRT; 696 AA.	
AC	Q98T85		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Iuteinizing hormone receptor precursor.		
OS	Ictalurus punctatus (Channel catfish).		
OC	Euteleostomi; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;		
OC	Ictaluriidae; Ictalurus.		
NCBI_TaxID=7998;	[1]		
RN			

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21141975; PubMed=11207219;
RA Sampath Kumar R., Jijir S., Tran J.M.;
RT "Molecular Biology of Channel Catfish Gonadotropin Receptors: 1.
RT Cloning of a Functional luteinizing Hormone Receptor and Preovulatory
RT Induction of Gene Expression."
RL Biol. Reprod. 64:1010-1018(2001).
DR EMBL; AF285181; AAK16066.1; -.
DR HSSP; P22889; IUTV.
DR InterPro; IPR000276; GPCR_Rhodpn.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; UNKNOWN_1.
KM Receptor; Signal.
FT SIGNAL.
SQ SEQUENCE 696 AA; 77822 MW; 4A1FB19CDE070AB6 CRC64;

Query Match 17.4%; Score 143.5; DB 13; Length 696;
Best Local Similarity 29.5%; Pred. No. 3.3e-06;
Matches 49; Conservative 30; Mismatches 66; Indels 21; Gaps 6;

QY 1 AOIYSVAIFLGINLAFFIIVFSYSGSMFYHQSATATATIRNQVKEMILAKRFFPIVF 60
DB 527 SQGVVAV-LVINVAALFVWCSYAGIYLSVRNPNVPTRRGHR-----MAKGMALIF 579
QY 61 TDALCWIPF-FVAKPLSLQVEIPGTTISWVIVGYSAINSALNPILYTLTRPFKE---- 115
DB 580 TDFLCMAPISFALISALHMPILSVSQSKILLIFYPINSLCPHFLYITPFAFRDMCL 639
QY 116 MHRFMVNRKSKMSDKGI-----RKMLH--HSSGCKGCR 152
DB 640 LLSKCGCHQAERYRQGLGFLTKARKKRVKPKSKNFRAYHVK 685

RESULT 7
Q64183 PRELIMINARY; PRT; 688 AA.
AC Q64183;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 01, Last annotation update)
DE Follicle-stimulating hormone receptor.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10118;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92149579; PubMed=1738373;
RA Heckert L.L., Daley J.J., Griswold M.D.;
RT "Structural organization of the follicle-stimulating hormone receptor
RT gene."
RL Mol. Endocrinol. 6:70-80(1992).
DR EMBL; S81198; AAB21415.2; -.
DR EMBL; S81117; AAB21415.2; JOINED.
DR EMBL; S81119; AAB21415.2; JOINED.
DR EMBL; S81121; AAB21415.2; JOINED.
DR EMBL; S81171; AAB21415.2; JOINED.
DR EMBL; S81174; AAB21415.2; JOINED.
DR EMBL; S81178; AAB21415.2; JOINED.
DR EMBL; S81183; AAB21415.2; JOINED.
DR EMBL; S81185; AAB21415.2; JOINED.
DR EMBL; S81194; AAB21415.2; JOINED.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_3.
DR Pfam; PF01462; LRRNT; 1.

```

```

DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
KM Receptor.
SQ SEQUENCE 688 AA; 77341 MW; 441F0D9EFD01DF18 CRC64;

Query Match 16.9%; Score 139.5; DB 11; Length 688;
Best Local Similarity 27.3%; Pred. No. 8.3e-06;
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;

QY 1 AOIYSVAIFLGINLAFFIIVFSYSGSMFYHQSATATATIRNQVKEMILAKRFFPIVF 60
DB 522 SQLYVVAL-LVINVAALFVIGCGCTHIYVIRNPITVS-----SSDTKIAKRWALIF 574
QY 61 TDALCWIPF-FVAKPLSLQVEIPGTTISWVIVGYSAINSALNPILYTLTRPFKE---- 115
DB 575 TDFLCMAPISFALISALKVLITVSKAKILLVLYPINSKANPFLVAFKTRFRDFFI 634
QY 116 MHRF-----WHYRQKMSDKGIRK-----MLHSS 144
DB 635 LLSKCGCEMQAIVRTETSSATNHFARKSHGSAFRTNSVYLVLPHSS 686

RESULT 8
Q90XC7 PRELIMINARY; PRT; 307 AA.
AC Q90XC7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Purative thyrotropin receptor TSHR (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Actinopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Yaderie F., Male R.;
RT "Cloning and characterization of Atlantic salmon ovarian thyrotropin
RT receptor (TSHR) like cDNA."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF406603; AAK98600.1; -.
DR InterPro; IPR000276; GPCR_Rhodpn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
KM Receptor.
FT NON TER.
SQ SEQUENCE 307 AA; 34289 MW; 0E3B8032DAF68470 CRC64;

Query Match 16.7%; Score 138; DB 13; Length 307;
Best Local Similarity 28.9%; Pred. No. 5e-06;
Matches 35; Conservative 25; Mismatches 41; Indels 20; Gaps 4;

QY 1 AOIYSVAIFLGINLAFFIIVFSYSGSMFYHQSATATATIRNQVKEMILAKRFFPIVF 60
DB 115 AOIYSVAL-LVINVAALFVIGCGCTHIYVIRNPITVS-----SKDTIAKRWALIF 167
QY 61 TDALCWIPF-FVAKPLSLQVEIPGTTISWVIVGYSAINSALNPILYTLTRPFKE---- 113
DB 168 TDFLCMAPISFYMSPVVDPL-----ITVNSKILLVLYPINSKANPFLVAFITKAF 221
QY 114 K 114
DB 222 R 222

RESULT 9
Q91949 PRELIMINARY; PRT; 814 AA.
AC Q91949;

```



DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Thyrotropin receptor A.  
 GN TSH-RA.  
 OS Oncorhynchus rhodurus (amago).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=41164;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20462948; PubMed=11006115;  
 RA Oba Y., Hira T., Yoshida Y., Kobayashi T., Nagahama Y.;  
 RT "Cloning, Functional Characterization, and Expression of Thyrotropin  
 Receptor in the Thyroid of Amago Salmon (Oncorhynchus rhodurus).";  
 RL Biochem. Biophys. Res. Commun. 276:258-263 (2000).  
 DR EMBL; AB030954; BAB07800.1; -.  
 DR HSSP; P16473; 1XUM.  
 DR InterPro; IPR000276; GPCR\_Rhodpn.  
 DR InterPro; IPR001611; LRR.  
 DR Pfam; PF000001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR Receptor.  
 SQ SEQUENCE 814 AA; 91091 MW; 1FC9A2AB8DC911E CRC64;

Query Match 16.2%; Score 134; DB 13; Length 814;  
 Best Local Similarity 26.8%; Pred. No. 3.6e-05;  
 Matches 44; Conservative 32; Mismatches 60; Indels 28; Gaps 6;

QY 1 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 DB 624 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 624 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 QY 61 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114  
 DB 677 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114  
 677 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114  
 QY 115 -----EMIRFHWNYRQRKMSDKIRKMLHSSGKCGCHRR 153  
 DB 737 LLSKVGLCQRRALFR-GQTVSSKG-----SSGVCHQGR 770  
 737 LLSKVGLCQRRALFR-GQTVSSKG-----SSGVCHQGR 770

## RESULT 10

QY 09DE63 PRELIMINARY; PRT; 424 AA.  
 DB 09DE63;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE GnRH receptor-3 (Type III GnRH).  
 GN BFGNRH-3.  
 OS Rana catesbeiana (bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranae; Rana.  
 NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2106155; PubMed=11120886;  
 RA Wang L., Bogerd J., Choi H.S., Seong J.Y., Son J.M., Chun S.Y.,  
 RA Blomentröth M., Tioekle B.E., Millar R.P., Yu W.H., McCann S.M.,  
 RA Kwon H.B.;  
 RT "Three distinct types of GnRH receptor characterized in the  
 bullfrog.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:361-366 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21407809; PubMed=11517181;  
 RA Wang L., Oh D.Y., Bogerd J., Choi H.S., Ahn R.S., Seong J.Y.,

RA Kwon H.B.;  
 RT "Inhibitory activity of alternative splice variants of the bullfrog  
 GnRH receptor-3 on wild-type receptor signaling.";  
 RL Endocrinology 142:4015-4025 (2001).  
 DR EMBL; AF144062; AAG42574.1; -.  
 DR EMBL; AF224277; AAL11631.1; -.  
 DR InterPro; IPR000276; GPCR\_Rhodpn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR Receptor.  
 KW SEQUENCE 424 AA; 47613 MW; E771F7403AB44AD6 CRC64;

Query Match 16.2%; Score 133.5; DB 13; Length 424;  
 Best Local Similarity 22.6%; Pred. No. 2e-05;  
 Matches 44; Conservative 33; Mismatches 73; Indels 45; Gaps 6;

QY 3 IYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 52  
 DB 224 IYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 52  
 224 IYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 52  
 QY 53 KRFFIVFTDALCWIPFIVA-----KPLSLQVEIPGTITSWVIGYSAINSALNPILY 106  
 DB 284 KRFFIVFTDALCWIPFIVA-----KPLSLQVEIPGTITSWVIGYSAINSALNPILY 106  
 284 KRFFIVFTDALCWIPFIVA-----KPLSLQVEIPGTITSWVIGYSAINSALNPILY 106  
 QY 107 TLTPFPKEMIRFHWNYRQRKMSDKIRKMLHSSGKCGCHRR 145  
 DB 343 GLFTHFREIRFHWNYRQRKMSDKIRKMLHSSGKCGCHRR 145  
 343 GLFTHFREIRFHWNYRQRKMSDKIRKMLHSSGKCGCHRR 145  
 QY 146 -----GKCHGR 153  
 DB 403 TGVLHSGKCHGR 417  
 403 TGVLHSGKCHGR 417

## RESULT 11

QY 08SPS8 PRELIMINARY; PRT; 701 AA.  
 DB 08SPS8;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Luteinizing hormone receptor.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=HOLSTEIN; TISSUE=CORPUS LUTEUM;  
 RA Kawate N., Tamada H., Inaba T., Sawada T.;  
 RT "Molecular cloning of a full-length cDNA encoding bovine luteinizing  
 hormone receptor and its expression in COS-7 cells.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF491303; AAM09535.1; -.  
 KW Receptor.  
 SQ SEQUENCE 701 AA; 78455 MW; D70AB62B265CCE CRC64;

Query Match 16.0%; Score 132; DB 6; Length 701;  
 Best Local Similarity 29.6%; Pred. No. 4.9e-05;  
 Matches 34; Conservative 26; Mismatches 47; Indels 8; Gaps 3;

QY 1 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 DB 526 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 526 AQTYSVAIFLGINLAFLIIVSYSGMSFYHQSATITETIRNOVKEMILAKRFFIVF 60  
 QY 61 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114  
 DB 579 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114  
 579 TDALCWIPF-FAKPLSLQVEIPGTITSWVIGYSAINSALNPILYTLTRPFK 114

## RESULT 12

QY 091948

```
ID Q01948 PRELIMINARY; PRT; 793 AA.
AC Q01948;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Thyrotropin receptor B.
GN STSH-RB.
OS Oncorhynchus rhodurus (amago).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=41164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20462948; PubMed=11006115;
RA Oba Y., Hirai T., Yoshizura Y., Kobayashi T., Nagahama Y.;
RT "Cloning, functional characterization, and expression of thyrotropin
RT Receptors in the thyroid of Amago Salmon (Oncorhynchus rhodurus).";
RL Biochem. Biophys. Res. Commun. 276:258-263(2000).
EMBL: AB030955; BAB07801.1; -.
HSSP: P16473; IXUM.
InterPro: IPR000276; GPCR_Rhodopsn.
InterPro: IPR001611; LRR.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR KEGG; K01948.
SQ SEQUENCE 793 AA; 89410 MW; 2842BF0D67ED93DC CRC64;

Query Match 16.0%; Score 132; DB 13; Length 793;
Best Local Similarity 27.3%; Pred. No. 5.6e-05;
Matches 33; Conservative 26; Mismatches 40; Indels 20; Gaps 4;

QY 1 AQIYVAIFLGINLAFLIIVSYGSMFYVHQAATATETIRNVKKEMLAKRFFPIVF 60
DB 598 AQQYIISVLI-LNIIAFLVICCYKVIKCAVHNPYCSG-----SKDTIARKEMLILF 650
QY 61 TDALCWIPF-----FVAKPLSLQVEIPGTITSWVWIGVSAINSALNPILYLTTRPK 113
DB 651 TDFTCMAPISFYAMSAVVDRL-----ITVNSKILLVLFYPLNSCANPFLVAIFTKAF 704
QY 114 K 114
DB 705 R 705

UT 13
96
ID Q15996 PRELIMINARY; PRT; 699 AA.
AC Q15996;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Luteinizing hormone receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92017881; PubMed=1922095;
RA Uta X.C., Okawa M., Bo M., Tanaka T., Ny T., Boime I., Hsueh A.J.;
RT "Expression of human luteinizing hormone (LH) receptor: interaction
RT with LH and chorionic gonadotropin from human but not equine, rat, and
RT ovine species.";
RL Mol. Endocrinol. 5:759-768(1991).
EMBL: S57793; AAB19917.2; -.
HSSP: P22888; ILUT.
InterPro: IPR000276; GPCR_Rhodopsn.
InterPro: IPR001611; LRR.
```

```
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR KEGG; K01948.
SQ SEQUENCE 699 AA; 78642 MW; 2E3D93F4621BA842 CRC64;

Query Match 15.9%; Score 131; DB 4; Length 699;
Best Local Similarity 29.6%; Pred. No. 6.2e-05;
Matches 34; Conservative 25; Mismatches 48; Indels 8; Gaps 3;

QY 1 AQIYVAIFLGINLAFLIIVSYGSMFYVHQAATATETIRNVKKEMLAKRFFPIVF 60
DB 524 SOVYITITLI-LNVVAFFLICCYKIIFAVRNPELMT-----NKDTKIAKKMALILF 576
QY 61 TDALCWIPF-----FVAKPLSLQVEIPGTITSWVWIGVSAINSALNPILYLTTRPK 114
DB 577 TDFTCMAPISFYAMSAVVDRL-----ITVNSKILLVLFYPLNSCANPFLVAIFTKAF 631

RESULT 14
ID Q14751 PRELIMINARY; PRT; 701 AA.
AC Q14751;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE LUTEINISING hormone-choriogonadotropin receptor (luteinizing hormone
DE receptor).
GN LHGR OR LHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96023956; PubMed=7556872;
RA Atger M., Mirrahi M., Sar S., Leflem L., Dessen P., Milgrom E.;
RT "Structure of the human luteinizing hormone-choriogonadotropin
RT receptor gene: unusual promoter and 5' non-coding regions.";
RL Mol. Cell. Endocrinol. 111:113-123(1995).
RN [2]
RP SEQUENCE OF 1-56 FROM N.A.
RA Tsai-Morris C.H., Geng Y., Buczek E., Dufau M.L.;
RT "Human luteinizing hormone receptor gene containing 6 bp insertion in
RT exon 1 coding region with identical 5' flanking sequence as LHR2 exon
RT 1 coding region.";
RL Hum. Hered. 49:0-0(1999).
DR EMBL; X84753; CAAS9234.1; -.
DR EMBL; X84754; CAAS9234.1; JOINED.
DR EMBL; X84755; CAAS9234.1; JOINED.
DR EMBL; X84756; CAAS9234.1; JOINED.
DR EMBL; X84757; CAAS9234.1; JOINED.
DR EMBL; X84758; CAAS9234.1; JOINED.
DR EMBL; X84759; CAAS9234.1; JOINED.
DR EMBL; X84760; CAAS9234.1; JOINED.
DR EMBL; X84761; CAAS9234.1; JOINED.
DR EMBL; X84762; CAAS9234.1; JOINED.
DR EMBL; X84763; CAAS9234.1; JOINED.
DR EMBL; AF082076; AAC98291.1; -.
HSSP: P22888; ILUT.
InterPro: IPR000276; GPCR_Rhodopsn.
InterPro: IPR001611; LRR.
InterPro: IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR KEGG; K01948.
SQ SEQUENCE 701 AA; 78757 MW; E69FF5BABC409D4 CRC64;
```

Query Match 15.9%; Score 131; DB 4; Length 701;  
 Best Local Similarity 29.6%; Pred. No. 6.2e-05;  
 Matches 34; Conservative 25; Mismatches 48; Indels 8; Gaps 3;

QY 1 AQIYSVAIFGINIAFIIVFSYSGMSFYSVHOSAITATEIRNOVKEMILAKRFFFIIV 60  
 DB 526 SQVYILITLI-LNVVAFPIICACYIKIYFAVRNPELMAT-----NDQYIAKMAVLIF 578  
 QY 61 TDALCWIPD-FVAKPLSLQVEIPGTTISWVIGYSAINSALNPILYTLTRPDK 114  
 DB 579 TDFTCMAPISFSAFAKXVPLITVINSKVLVLFYFINSKANPFLYAITFTKTFQ 633

## RESULT 15

Q91BN7 PRELIMINARY; PRT; 779 AA.  
 ID 091BN7  
 AC 091BN7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 Ds Thyrotropin receptor precursor.  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 OC NCBI\_TaxID=34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVARY;  
 RX MEDLINE=20457109; PubMed=1100515;  
 RA Kumar R.S., Ijiri S., Knight K., Swanson P., Dittman A., Alok D.,  
 RA Zohar Y., Trant J.M.;  
 RT "Cloning and functional expression of a thyrotropin receptor from the  
 RT gonads of a vertebrate (bony fish): potential thyroid-independent role  
 RT for thyrotropin in reproduction."  
 RL Mol. Cell. Endocrinol. 167:1-9(2000).  
 DR EMBL; AF239761; AAF80596.1; -.  
 DR HSSP; P16473; 1XUM.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR InterPro; IPR001611; LRR.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 2.  
 DR PRINTS; PR00237; GPCR\_RHODOPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 DR Receptor; Signal.  
 DR SIGNAL 1 22 POTENTIAL.  
 SEQUENCE 779 AA; 87400 MW; 8C86D121B676A9EF CRC64;

Query Match 15.9%; Score 131; DB 13; Length 779;  
 Best Local Similarity 29.6%; Pred. No. 6.9e-05;  
 Matches 34; Conservative 27; Mismatches 46; Indels 8; Gaps 3;

QY 1 AQIYSVAIFGINIAFIIVFSYSGMSFYSVHOSAITATEIRNOVKEMILAKRFFFIIV 60  
 DB 588 AQYIILISV-LVNLILAVVICACFKIYCAVHNPHRSG-----SKDTNIAKMAVLIF 640  
 QY 61 TDALCWIPD-FVAKPLSLQVEIPGTTISWVIGYSAINSALNPILYTLTRPDK 114  
 DB 641 TDFTCMAPISFSAFAKXVPLITVINSKVLVLFYFINSKANPFLYAITFTKTFQ 695

Search completed: February 3, 2003, 14:24:43  
 Job time : 75 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:19:02 ; Search time 22 Seconds  
(without alignments)  
209.973 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 826

Sequence: 1 AQLYSVAIFLGINLAFAII.....HMLHSSGKCGHCRCHLS 157

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCUTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	147.5	17.9	792	4	US-07-741-453A-56 Sequence 56, Appl
2	140.5	17.0	696	4	US-07-757-342D-4 Sequence 4, Appl
3	139.5	16.9	682	4	US-07-757-342D-6 Sequence 6, Appl
4	137.5	16.6	795	4	US-07-741-453A-55 Sequence 55, Appl
5	136	16.5	370	1	US-08-118-270-18 Sequence 18, Appl
6	136	16.5	700	5	PCT-US93-08528-18 Sequence 18, Appl
7	134	16.2	700	4	US-07-757-342D-3 Sequence 3, Appl
8	131	15.9	611	4	US-07-757-342D-8 Sequence 8, Appl
9	131	15.9	636	4	US-07-757-342D-7 Sequence 7, Appl
10	131	15.9	674	4	US-07-757-342D-10 Sequence 10, Appl
11	131	15.9	699	4	US-07-757-342D-2 Sequence 2, Appl
12	125.5	15.2	420	4	US-08-795-876-33 Sequence 33, Appl
13	125.5	15.2	436	4	US-08-795-876-38 Sequence 38, Appl
14	125.5	15.2	436	4	US-08-795-876-38 Sequence 38, Appl
15	125.5	15.2	695	1	US-08-487-886-2 Sequence 2, Appl
16	125.5	15.2	695	3	US-08-487-886-2 Sequence 2, Appl
17	125.5	15.2	695	3	US-08-487-886-2 Sequence 2, Appl
18	124	15.0	338	3	US-08-988-876-8 Sequence 8, Appl
19	124	15.0	338	3	US-08-988-876-8 Sequence 8, Appl
20	123.5	15.0	325	2	US-08-467-947A-29 Sequence 29, Appl
21	123.5	15.0	325	2	US-08-467-947A-29 Sequence 29, Appl
22	122	14.8	764	4	US-07-741-453A-54 Sequence 54, Appl
23	120.5	14.6	764	4	US-07-741-453A-54 Sequence 54, Appl
24	119.5	14.5	332	4	US-08-293-563-5 Sequence 5, Appl
25	119.5	14.5	378	4	US-09-082-088-2 Sequence 2, Appl
26	119.5	14.5	378	4	US-09-082-088-2 Sequence 2, Appl
27	119.5	14.5	517	2	US-08-467-568-10 Sequence 10, Appl

28	119.5	14.5	517	2	US-09-030-582-10 Sequence 10, Appl
29	118.5	14.3	515	1	US-08-194-338-3 Sequence 3, Appl
30	118	14.3	764	4	US-07-741-453A-61 Sequence 61, Appl
31	117.5	14.2	336	4	US-08-118-270-54 Sequence 54, Appl
32	117.5	14.2	336	5	PCT-US93-08528-54 Sequence 54, Appl
33	117.5	14.2	420	4	US-09-255-368-6 Sequence 6, Appl
34	117.5	14.2	513	2	US-08-406-855A-21 Sequence 21, Appl
35	117.5	14.2	513	3	US-09-206-899-21 Sequence 21, Appl
36	117.5	14.2	515	1	US-08-444-734A-7 Sequence 7, Appl
37	117.5	14.2	515	1	US-08-722-001-25 Sequence 25, Appl
38	117.5	14.2	515	2	US-08-406-855A-22 Sequence 22, Appl
39	117.5	14.2	515	3	US-09-206-899-22 Sequence 22, Appl
40	117.5	14.2	515	4	US-09-032-742-5 Sequence 5, Appl
41	117.5	14.2	515	4	US-09-688-415-9 Sequence 9, Appl
42	117.5	14.2	515	4	US-09-688-415-10 Sequence 10, Appl
43	117.5	14.2	764	4	US-07-757-342D-5 Sequence 5, Appl
44	117.5	14.2	764	4	US-07-741-453A-59 Sequence 59, Appl
45	117	14.2	332	1	US-08-349-696-23 Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-07-741-453A-56  
; Sequence 56, Application US/07741453A  
; Patent No. 6228597  
GENERAL INFORMATION:  
APPLICANT: PARMENTER, MARC  
APPLICANT: LIBERT, FREDERIC  
APPLICANT: DUMONT, JACQUES  
APPLICANT: VASSART, GILBERT  
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR  
TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS  
TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSER: CUSHMAN DABY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: U.S.A.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/741,453A  
FILING DATE: 19911015  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 792 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-741-453A-56  
Query Match 17.9%; Score 147.5; DB 4; Length 792;  
Best Local Similarity 25.9%; Pred. No. 8.9e-08;  
Matches 42; Conservative 27; Mismatches 60; Indels 33; Gaps 4;  
QY 1 AQLYSVAIFL-----GINLAFAIIIVSYGSMFYSHQSAITATEIRNQVKEMILAKRF 55

Db 596 SQVYLITLPIGHCGINVAFLITICACIKITFAVONPELMAT-----NDOTKIAKKA 649  
QY 56 FFIIVFTDALCMIPFVAKPLSLQVE-----IPGTTISWVIGYSAINSLNPLIYTL 108  
Db 650 AVILFTDFTCMAPIGHGCGISFPAISALAKVPLITVNSKVLVLFYVNSCANPFLYAI 709  
QY 109 TTRPEKEMIRFMHNYRQRKSMDSKIRKMLHSSGCKGH 150  
Db 710 FTAKFR-----DPIGHGCFLLSKSGCKKH 736

RESULT 2  
US-07-757-342D-4  
; Sequence 4, Application US/07757342D  
; Patent No. 6218509  
; GENERAL INFORMATION:  
; APPLICANT: IGARASHI, Masao  
; MINGISHI, Takashi  
; NAKAMURA, Kazuto  
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &  
CUSHMAN  
STREET: 130 Water Street  
City: Boston  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,342D  
FILING DATE: 10-Sep-1991  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BUCKLEY, Linda M.  
REGISTRATION NUMBER: 31003  
REFERENCE/DOCKET NUMBER: 41226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)523-3400  
TELEFAX: (617)523-6440  
TELEX: 200291 STRE UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 696 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-07-757-342D-4

Query Match 17.0%; Score 140.5; DB 4; Length 696;  
Best Local Similarity 27.2%; Pred. No. 4.4e-07;  
Matches 41; Conservative 26; Mismatches 55; Indels 29; Gaps 4;  
QY 1 AOIYSVAIFLGINLAFFIIIVSYSGMSFYVHQSATITATEIRNOVKEMILAKRFFIIV 60  
Db 524 SQVYLITLPIGHCGINVAFLITICACIKITFAVONPELMAT-----NDOTKIAKKA 576  
QY 61 TDALCMIPFVAKPLSLQVEIPGTTISWVIGYSAINSLNPLIYTLTTRPEKEMIR 119  
Db 577 TDFTCMAPISFPAISALAKVPLITVNSKVLVLFYVNSCANPFLYAIPTKAFR----- 631  
QY 120 FMHNYRQRKSMDSKIRKMLHSSGCKGH 150  
Db 632 -----RDFLLSKSGCKKH 646

RESULT 3  
US-07-757-342D-6  
; Sequence 6, Application US/07757342D  
; Patent No. 6218509  
; GENERAL INFORMATION:  
; APPLICANT: IGARASHI, Masao  
; MINGISHI, Takashi  
; NAKAMURA, Kazuto  
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &  
CUSHMAN  
STREET: 130 Water Street  
City: Boston  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/757,342D  
FILING DATE: 10-Sep-1991  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BUCKLEY, Linda M.  
REGISTRATION NUMBER: 31003  
REFERENCE/DOCKET NUMBER: 41226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)523-3400  
TELEFAX: (617)523-6440  
TELEX: 200291 STRE UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 692 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-07-757-342D-6

Query Match 16.9%; Score 139.5; DB 4; Length 692;  
Best Local Similarity 27.3%; Pred. No. 5.6e-07;  
Matches 47; Conservative 30; Mismatches 60; Indels 35; Gaps 6;  
QY 1 AOIYSVAIFLGINLAFFIIIVSYSGMSFYVHQSATITATEIRNOVKEMILAKRFFIIV 60  
Db 526 SQVYLITLPIGHCGINVAFLITICACIKITFAVONPELMAT-----SDOTKIAKKA 578  
QY 61 TDALCMIPFVAKPLSLQVEIPGTTISWVIGYSAINSLNPLIYTLTTRPEKEMIR 115  
Db 579 TDFTCMAPISFPAISALAKVPLITVNSKVLVLFYVNSCANPFLYAIPTKAFRDF 638  
QY 116 MTHRF-----MHNYRQRKSMDSKIRK-----MLHSS 144  
Db 639 LSKFCYEMQAOIYRTETSSATNHFARKSKSSAPVTVNSYVLPINHSS 690

RESULT 4  
US-07-741-453A-55  
; Sequence 55, Application US/07741453A  
; GENERAL INFORMATION:  
; APPLICANT: PARMENTIER, MARC  
; APPLICANT: LIBERT, FREDERIC  
; APPLICANT: DUMONT, JACQUES  
; APPLICANT: VASSART, GILBERT  
TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR  
ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS  
AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES

NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CUSHMAN DABRY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: U.S.A.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/741,453A  
FILING DATE: 19911015  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: 91913/1107/US/ST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 795 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-741-453A-55

Query Match 16.6%; Score 137.5; DB 4; Length 795;  
Best Local Similarity 21.9%; Pred. No. 1.1e-06;  
Matches 43; Conservative 38; Mismatches 64; Indels 51; Gaps 7;

QY 1 AQIYVAFL-----GILAAFIIVSYSGMFSVHOSATATEIRNOVKEMILAKRF 55  
DB 599 SQVYLSILIRATHGCLVAVVACACIYIRIYFAVQNPBLTAP-----NKDKIKAKM 652  
QY 56 FIVFTDALCWPPIFVAKPLSLQVE-----IPGITTSWVIGYSAINSAALNPILYTL 108  
DB 653 ALITFTDCMARATHGCPISFPFSAFAKVPFLIVTMSKILLVLFYVNSCANPEFLYAI 712  
QY 109 TTRPKE-----MIRFW-----HNYROR-----SMDSKGIRK 137  
DB 713 FYKAFQDRATHGFLILSRFGCCRRALYRKEFSAYTSNCKNGFPGASKPSQATLK 772  
QY 138 HMLHSSGKCGHCR 153  
DB 773 --LSTYRATHGCHCOO 786

RESULT 5  
US-08-118-270-18  
Sequence 18, Application US/08118270  
Patent No. 5508384  
GENERAL INFORMATION:  
APPLICANT: Murphy, Randall B.  
APPLICANT: Schuster, David I.  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY=2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-18

Query Match 16.5%; Score 136; DB 1; Length 370;  
Best Local Similarity 28.5%; Pred. No. 6.3e-07;  
Matches 39; Conservative 24; Mismatches 50; Indels 24; Gaps 5;

QY 2 QIYVAIFLGINIAFIIVSYSGMFSV-----HOSATITETIRNOVKEMILAK 53  
DB 178 RYIVAKRTTKNLGAVMKEMS-NSKFTLRIMWSKNFEDTISTKAGHNPSSIAVK 236  
QY 54 REFF-----IVFDALCWPPIFVAKPLSLQVE-----IPGITTSWVIGYSAINSA 100  
DB 237 LFFPSSEKRAAKTLGIVGMLCWPPIFVAKPLSLQVE-----IPGITTSWVIGYSAINSA 294  
QY 101 LNPILYTLTTRPKEMI 117  
DB 295 LNPILYCSKPKRAL 311

RESULT 6  
PCT-US93-08528-18  
Sequence 18, Application PC/TUS9308528  
GENERAL INFORMATION:  
APPLICANT: New York University  
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 348  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08528  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT

```

1 /
2 / TELECOMMUNICATION INFORMATION:
3 / TELEPHONE: 202-628-5197
4 / TELEFAX: 202-737-3528
5 /
6 / TELEX: 248633
7 /
8 / INFORMATION FOR SEQ ID NO: 18:
9 /
10 / SEQUENCE CHARACTERISTICS:
11 / LENGTH: 370 amino acids
12 / TYPE: amino acid
13 / STRANDEDNESS: single
14 / TOPOLOGY: linear
15 /
16 / MOLECULE TYPE: peptide
17 /
18 / PCT-US93-08528-18
19 /
20 /
21 /
22 /
23 /
24 /
25 /
26 /
27 /
28 /
29 /
30 /
31 /
32 /
33 /
34 /
35 /
36 /
37 /
38 /
39 /
40 /
41 /
42 /
43 /
44 /
45 /
46 /
47 /
48 /
49 /
50 /
51 /
52 /
53 /
54 /
55 /
56 /
57 /
58 /
59 /
60 /
61 /
62 /
63 /
64 /
65 /
66 /
67 /
68 /
69 /
70 /
71 /
72 /
73 /
74 /
75 /
76 /
77 /
78 /
79 /
80 /
81 /
82 /
83 /
84 /
85 /
86 /
87 /
88 /
89 /
90 /
91 /
92 /
93 /
94 /
95 /
96 /
97 /
98 /
99 /
100 /
101 /
102 /
103 /
104 /
105 /
106 /
107 /
108 /
109 /
110 /
111 /
112 /
113 /
114 /
115 /
116 /
117 /
118 /
119 /
120 /
121 /
122 /
123 /
124 /
125 /
126 /
127 /
128 /
129 /
130 /
131 /
132 /
133 /
134 /
135 /
136 /
137 /
138 /
139 /
140 /
141 /
142 /
143 /
144 /
145 /
146 /
147 /
148 /
149 /
150 /
151 /
152 /
153 /
154 /
155 /
156 /
157 /
158 /
159 /
160 /
161 /
162 /
163 /
164 /
165 /
166 /
167 /
168 /
169 /
170 /
171 /
172 /
173 /
174 /
175 /
176 /
177 /
178 /
179 /
180 /
181 /
182 /
183 /
184 /
185 /
186 /
187 /
188 /
189 /
190 /
191 /
192 /
193 /
194 /
195 /
196 /
197 /
198 /
199 /
200 /
201 /
202 /
203 /
204 /
205 /
206 /
207 /
208 /
209 /
210 /
211 /
212 /
213 /
214 /
215 /
216 /
217 /
218 /
219 /
220 /
221 /
222 /
223 /
224 /
225 /
226 /
227 /
228 /
229 /
230 /
231 /
232 /
233 /
234 /
235 /
236 /
237 /
238 /
239 /
240 /
241 /
242 /
243 /
244 /
245 /
246 /
247 /
248 /
249 /
250 /
251 /
252 /
253 /
254 /
255 /
256 /
257 /
258 /
259 /
260 /
261 /
262 /
263 /
264 /
265 /
266 /
267 /
268 /
269 /
270 /
271 /
272 /
273 /
274 /
275 /
276 /
277 /
278 /
279 /
280 /
281 /
282 /
283 /
284 /
285 /
286 /
287 /
288 /
289 /
290 /
291 /
292 /
293 /
294 /
295 /
296 /
297 /
298 /
299 /
300 /
301 /
302 /
303 /
304 /
305 /
306 /
307 /
308 /
309 /
310 /
311 /
312 /
313 /
314 /
315 /
316 /
317 /
318 /
319 /
320 /
321 /
322 /
323 /
324 /
325 /
326 /
327 /
328 /
329 /
330 /
331 /
332 /
333 /
334 /
335 /
336 /
337 /
338 /
339 /
340 /
341 /
342 /
343 /
344 /
345 /
346 /
347 /
348 /
349 /
350 /
351 /
352 /
353 /
354 /
355 /
356 /
357 /
358 /
359 /
360 /
361 /
362 /
363 /
364 /
365 /
366 /
367 /
368 /
369 /
370 /
371 /
372 /
373 /
374 /
375 /
376 /
377 /
378 /
379 /
380 /
381 /
382 /
383 /
384 /
385 /
386 /
387 /
388 /
389 /
390 /
391 /
392 /
393 /
394 /
395 /
396 /
397 /
398 /
399 /
400 /
401 /
402 /
403 /
404 /
405 /
406 /
407 /
408 /
409 /
410 /
411 /
412 /
413 /
414 /
415 /
416 /
417 /
418 /
419 /
420 /
421 /
422 /
423 /
424 /
425 /
426 /
427 /
428 /
429 /
430 /
431 /
432 /
433 /
434 /
435 /
436 /
437 /
438 /
439 /
440 /
441 /
442 /
443 /
444 /
445 /
446 /
447 /
448 /
449 /
450 /
451 /
452 /
453 /
454 /
455 /
456 /
457 /
458 /
459 /
460 /
461 /
462 /
463 /
464 /
465 /
466 /
467 /
468 /
469 /
470 /
471 /
472 /
473 /
474 /
475 /
476 /
477 /
478 /
479 /
480 /
481 /
482 /
483 /
484 /
485 /
486 /
487 /
488 /
489 /
490 /
491 /
492 /
493 /
494 /
495 /
496 /
497 /
498 /
499 /
500 /
501 /
502 /
503 /
504 /
505 /
506 /
507 /
508 /
509 /
510 /
511 /
512 /
513 /
514 /
515 /
516 /
517 /
518 /
519 /
520 /
521 /
522 /
523 /
524 /
525 /
526 /
527 /
528 /
529 /
530 /
531 /
532 /
533 /
534 /
535 /
536 /
537 /
538 /
539 /
540 /
541 /
542 /
543 /
544 /
545 /
546 /
547 /
548 /
549 /
550 /
551 /
552 /
553 /
554 /
555 /
556 /
557 /
558 /
559 /
560 /
561 /
562 /
563 /
564 /
565 /
566 /
567 /
568 /
569 /
570 /
571 /
572 /
573 /
574 /
575 /
576 /
577 /
578 /
579 /
580 /
581 /
582 /
583 /
584 /
585 /
586 /
587 /
588 /
589 /
590 /
591 /
592 /
593 /
594 /
595 /
596 /
597 /
598 /
599 /
600 /
601 /
602 /
603 /
604 /
605 /
606 /
607 /
608 /
609 /
610 /
611 /
612 /
613 /
614 /
615 /
616 /
617 /
618 /
619 /
620 /
621 /
622 /
623 /
624 /
625 /
626 /
627 /
628 /
629 /
630 /
631 /
632 /
633 /
634 /
635 /
636 /
637 /
638 /
639 /
640 /
641 /
642 /
643 /
644 /
645 /
646 /
647 /
648 /
649 /
650 /
651 /
652 /
653 /
654 /
655 /
656 /
657 /
658 /
659 /
660 /
661 /
662 /
663 /
664 /
665 /
666 /
667 /
668 /
669 /
670 /
671 /
672 /
673 /
674 /
675 /
676 /
677 /
678 /
679 /
680 /
681 /
682 /
683 /
684 /
685 /
686 /
687 /
688 /
689 /
690 /
691 /
692 /
693 /
694 /
695 /
696 /
697 /
698 /
699 /
700 /
701 /
702 /
703 /
704 /
705 /
706 /
707 /
708 /
709 /
710 /
711 /
712 /
713 /
714 /
715 /
716 /
717 /
718 /
719 /
720 /
721 /
722 /
723 /
724 /
725 /
726 /
727 /
728 /
729 /
730 /
731 /
732 /
733 /
734 /
735 /
736 /
737 /
738 /
739 /
740 /
741 /
742 /
743 /
744 /
745 /
746 /
747 /
748 /
749 /
750 /
751 /
752 /
753 /
754 /
755 /
756 /
757 /
758 /
759 /
760 /
761 /
762 /
763 /
764 /
765 /
766 /
767 /
768 /
769 /
770 /
771 /
772 /
773 /
774 /
775 /
776 /
777 /
778 /
779 /
780 /
781 /
782 /
783 /
784 /
785 /
786 /
787 /
788 /
789 /
790 /
791 /
792 /
793 /
794 /
795 /
796 /
797 /
798 /
799 /
800 /
801 /
802 /
803 /
804 /
805 /
806 /
807 /
808 /
809 /
810 /
811 /
812 /
813 /
814 /
815 /
816 /
817 /
818 /
819 /

```

```

Query Match 16.2%; Score 134; DB 4; Length 700;
Best Local Similarity 27.2%; Pred. No. 2.2e-06;
Matches 34; Conservative 31; Mismatches 48; Indels 12; Gaps 4

QY 1 AQIYSAIFLGINLAAPITIVFSGSMFVSVHOSATATATEIRNOVKEMILAKRFFIIVF 60
DB 528 SOYIILISILIL-LNVVAFVVICACINIIYFVNOQBELTAP-----NKDRIKAKKAILIF 580
QY 61 TDAICWPII-FVAKPSLSLOVEIPGTTSMWVIGSAINSAINPIIYTLTRPPEK---- 115
DB 581 TDTFCMAPISFFAISAFAKVPILTVNSKILLVLFPVNSCANPFLYAIFTAFOQDFLL 640
QY 116 MIHF 120
DB 641 LLSRF 645

RESULT 8
US-07-757-342D-8
; Sequence 8, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; MINEGISHI, Takashi
; NAKAMURA, Kazuo
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 611 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-07-757-342D-8

Query Match 15.9%; Score 131; DB 4; Length 611;
Best Local Similarity 29.6%; Pred. No. 4e-06;
Matches 34; Conservative 25; Mismatches 48; Indels 8; Gaps 3;

QY 1 AQIYSAIFLGINLAAPITIVFSGSMFVSVHOSATATATEIRNOVKEMILAKRFFIIVF 60
DB 436 SOYIILISILIL-LNVVAFVVICACINIIYFVNOQBELTAP-----NKDRIKAKKAILIF 488
QY 61 TDAICWPII-FVAKPSLSLOVEIPGTTSMWVIGSAINSAINPIIYTLTRPPEK 114
DB 489 TDTFCMAPISFFAISAFAKVPILTVNSKILLVLFPVNSCANPFLYAIFTAFTKFO 543

```







```

RESULT 14
US-08-795-876-2
; Sequence 2, Application US/08795876
; Patent No. 6403505
; GENERAL INFORMATION:
; APPLICANT: Gershengorn, Marvin C.
; APPLICANT: Geras-Raaka, Elizabeth
; APPLICANT: Nussenzweig, Daniel R.
; TITLE OF INVENTION: STRATEGY TO CLONE DRUGS FOR G PROTEIN
; COUPLED RECEPTORS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,876
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BRAVAN, SUSAN J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/1280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-795-876-2
;
Query Match          15 2%; Score 125.5; DB 4; Length 436;
Best Local Similarity 26.1%; Pred. No. 1.le-05;
Matches 43; Conservative 30; Mismatches 59; Indels 33; Gaps

      1 AQLISVAIFLGINAAFLIIVFSYGSMEYSVHQSATATEIRNOVKKEMILAKRPFIVE 60
        :|:::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    268 SQLYVMSS-LVLNVLAFVVICGCIHIYLYTRNNIYSS-----SSDTRIARKMALIF 3200
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   61 TDAICWPI-FVAKPRLSLLOVELPGTITSWVGIVSAINSALNDILYTLTPRK----- 1144
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    321 TDPLCMAPISFFAISAISIKVPLIVSKAKILLVLFHPINSCANFLYAIFTKRRDRDFI 3800
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||

QY   115 -----EMTHRFMHNYRQKKSDSGKIRGHMLMHSGGCGCGR 151
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    381 LLSKCGCYEMQAQT--YRTETSTV-----HNHTPRNGHC 413
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 15
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 5744448
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweichardt, Rene Lynn
; APPLICANT: Cheng, Shirley Vui Yen
; APPLICANT: Nugent, No. 5744448een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; HORMONE
; NUMBER OF SEQUENCES: 2

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSER: Area-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Masell via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 14:21:18 ; Search time 17 Seconds  
(without alignments)  
186.355 Million cell updates/sec

Title: US-10-049-568-2

Perfect score: 856

Sequence: 1 AQLYSVAIFLGINIAAFIIL.....HMLHSSGKCHCRCHLS 157

Scoring table: BLOSUM62

Searched: 122226 seqs, 20178551 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep:\*

9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*

10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*

11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*

12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*

14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	77.2	396	10	US-09-895-686-4
2	638	77.2	757	10	US-09-928-175-24
3	632	76.5	176	9	US-09-989-442-105
4	632	76.5	176	10	US-09-764-853-557
5	632	76.5	176	10	US-09-764-877-2009
6	409	49.5	718	10	US-09-928-175-21
7	409	49.5	718	10	US-09-928-175-20
8	392.5	47.5	646	10	US-09-928-175-13
9	392.5	47.5	682	10	US-09-928-175-12
10	392.5	47.5	694	10	US-09-928-175-8
11	392.5	47.5	718	10	US-09-928-175-3
12	392.5	47.5	730	10	US-09-928-175-7
13	392.5	47.5	754	10	US-09-928-175-2
14	139.5	16.9	675	10	US-09-877-804-7
15	139.5	16.9	692	10	US-09-877-804-6
16	134	16.2	674	10	US-09-877-804-3
17	134	16.2	700	10	US-09-877-804-2
18	131	15.9	458	10	US-09-862-7674-5
19	131	15.9	699	10	US-09-804-626-6

20	130	15.7	861	10	US-09-804-5518-20	Sequence 20, Appl
21	125.5	15.2	658	10	US-09-862-7674-9	Sequence 9, Appl
22	125.5	15.2	695	10	US-09-804-626-8	Sequence 8, Appl
23	124	15.0	538	10	US-09-919-497-77	Sequence 77, Appl
24	121	14.6	764	9	US-10-045-6248-2	Sequence 2, Appl
25	119.5	14.5	378	9	US-09-971-228-7	Sequence 7, Appl
26	119.5	14.5	378	10	US-09-842-316-7	Sequence 7, Appl
27	119.5	14.5	378	10	US-09-731-030A-18	Sequence 18, Appl
28	119.5	14.5	378	12	US-10-037-616-21	Sequence 21, Appl
29	119.5	14.5	398	10	US-09-815-333-2	Sequence 2, Appl
30	119.5	14.5	398	10	US-09-842-316-2	Sequence 2, Appl
31	119.5	14.5	398	12	US-10-037-616-26	Sequence 26, Appl
32	119.5	14.5	517	10	US-09-951-622-10	Sequence 10, Appl
33	118.5	14.3	338	10	US-09-826-008-20	Sequence 20, Appl
34	117.5	14.2	332	9	US-09-983-442-127	Sequence 127, App
35	117.5	14.2	332	10	US-09-764-853-660	Sequence 660, App
36	117.5	14.2	420	9	US-09-866-248A-6	Sequence 6, Appl
37	117.5	14.2	497	12	US-10-052-589-2	Sequence 2, Appl
38	113.5	13.7	400	10	US-09-966-871-85	Sequence 85, Appl
39	113.5	13.7	400	12	US-10-039-645-85	Sequence 85, Appl
40	113.5	13.7	415	10	US-09-823-114-20	Sequence 20, Appl
41	112.5	13.6	359	10	US-09-862-767A-7	Sequence 7, Appl
42	112.5	13.6	400	10	US-09-862-871-78	Sequence 78, Appl
43	112.5	13.6	400	12	US-10-039-645-78	Sequence 78, Appl
44	112	13.6	400	9	US-09-971-228-12	Sequence 12, Appl
45	112	13.6	400	10	US-09-812-272-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1

US-09-895-686-4

Sequence 4, Application US/09895686

Patent No. US2002010655A1

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Lal, Preeti

APPLICANT: Tang, Y. Tom

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: HUMAN GPCR PROTEINS

FILE REFERENCE: PC-0044 CIP

CURRENT APPLICATION NUMBER: US/09/895.686

CURRENT FILING DATE: 2001-06-28

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 396

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US2002010655A1 2488822CD1

US-09-895-686-4

Query Match 77.2% Score 638; DB 10; Length 396;  
Best Local Similarity 94.8%; Pred. No. 5.1e-62;  
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AQLYSVAIFLGINIAAFIILVFVSYSMFYSVHQSATITATEIRNOVKKEMILAKFFFIIVF 60

DB 215 AQLYSVAIFLGINIAAFIILVFVSYSMFYSVHQSATITATEIRNOVKKEMILAKFFFIIVF 274

QY 61 TPAICWIPFVAVKPSLQVEIRPGITISWVVTGSAITSANPLIYLTTPPFKEMIHFR 120

DB 275 TPAICWIPFVAVKPSLQVEIRPGITISWVVTGSAITSANPLIYLTTPPFKEMIHFR 334

QY 121 WHNYRQKSMDSKG 134

DB 335 WHNYRQKSMDSKG 348

RESULT 2

```
US-09-928-175-24
; Sequence 24, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
09-928-175-24

Query Match      77.2%; Score 638; DB 10; Length 757;
Best Local Similarity 94.8%; Pred. No. 1,1e-61;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AQIYSVAIFLGINLAFFIIVFSYGSMEYSVHGSAITATETIRNQVKKEMILAKRFFPIV 60
DB 576 AQIYSVAIFLGINLAFFIIVFSYGSMEYSVHGSAITATETIRNQVKKEMILAKRFFPIV 635
QY 61 TDALCWPIPIFAKPLSLQVETPGTITSMVYTGSAINSALNPILYTLTRPFKEMHRF 120
DB 636 TDALCWPIPIFAKPLSLQVETPGTITSMVYTGSAINSALNPILYTLTRPFKEMHRF 695
QY 121 WNNYRQRKSMDSKG 134
DB 696 WNNYRQRKSMDSKG 709

RESULT 3
US-09-989-442-105
; Sequence 105, Application US/09989442
; Patent No. US20030013649A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P208
; CURRENT APPLICATION NUMBER: US/09/989,442
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
```

```
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,239
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
```

PRIOR APPLICATION NUMBER: 60/240,960  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/239,935  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/239,937  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/241,787  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,474  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/246,532  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/249,216  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,210  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/226,681  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,759  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/225,213  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/227,182  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,214  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/235,836  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: 60/230,438  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/215,135  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: 60/225,266  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/249,218  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,208  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,213  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,212  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,207  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,245  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,244  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,217  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,211  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,215  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,264  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,214  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,297  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/231,242  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,240  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,081  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,080  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,414  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,244  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/233,064

PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/233,063  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,397  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,399  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,401  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/241,808  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,826  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,786  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,221  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,475  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/231,243  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/233,065  
 PRIOR FILING DATE: 2000-09-14

Query Match 76.5%; Score 632; DB 9; Length 176;  
 Best Local Similarity 94.0%; Pred. No. 8,4e-62;  
 Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AOIYSVAIFLGINLAFFIIIVSYGSMFYSHQSAITATEIRNOVKKEMIIAKRFFIIVF 60  
 DB 1 AOIYSVAIFLGINLAFFIIIVSYGSMFYSHQSAITATEIRNOVKKEMIIAKRFFIIVF 60  
 QY 61 TDALCWPILFVVKPLSLQVEIRPGITTSWVYIGYAINSAINPILYTLTPPFKEMIRF 120  
 DB 61 TDALCWPILFVVKPLSLQVEIRPGITTSWVYIGYAINSAINPILYTLTPPFKEMIRF 120  
 QY 121 WHNYRQRKSMDSKG 134  
 DB 121 WHNYRQRKSMDSKG 134

RESULT 4  
 US-09-764-853-557  
 ; Sequence 557, Application US/09764853  
 ; Patent No. US20020090672A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P1206  
 ; CURRENT APPLICATION NUMBER: US/09/764,853  
 ; Prior application data removed - consult PAM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 939  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 557  
 ; LENGTH: 176  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (133)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-764-853-557

Query Match 76.5%; Score 632; DB 10; Length 176;  
 Best Local Similarity 94.0%; Pred. No. 8,4e-62;  
 Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 AOIYSVAIFLGINLAFFIIIVSYGSMFYSHQSAITATEIRNOVKKEMIIAKRFFIIVF 60  
 DB 1 AOIYSVAIFLGINLAFFIIIVSYGSMFYSHQSAITATEIRNOVKKEMIIAKRFFIIVF 60  
 QY 61 TDALCWPILFVVKPLSLQVEIRPGITTSWVYIGYAINSAINPILYTLTPPFKEMIRF 120

Db 61 TDALCWIPFVFLSLQVEIPGTTISWVVFILPINSALNPILYLTTRPFKEMTHRF 120  
QY 121 WHNYRORXSMDSKG 134  
Db 121 WHNYRORXSMDSKG 134

RESULT 5  
US-09-764-877-2009  
; Sequence 2009, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
EQ ID NO 2009  
LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: SITE  
; LOCATION: (133)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-877-2009

Query Match 76.5%; Score 632; DB 10; Length 176;  
Best Local Similarity 94.0%; Pred. No. 8.4e-62;  
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 1 AQISVAIFLGINLAFAITIVFSYSGMFYSVHOSAITATIRNOVKKEMILAKRFFPIVF 60  
Db 1 AQISVAIFLGINLAFAITIVFSYSGMFYSVHOSAITATIRNOVKKEMILAKRFFPIVF 60  
QY 61 TDALCWIPFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRF 120  
Db 61 TDALCWIPFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRF 120  
QY 121 WHNYRORXSMDSKG 134  
Db 121 WHNYRORXSMDSKG 134

ULT 6  
9-928-175-21  
; Sequence 21, Application US/09928175  
; Patent No. US20020123618A1  
; GENERAL INFORMATION:  
; APPLICANT: Paszty, Christopher J.  
; APPLICANT: Gong, Jianhua  
; APPLICANT: Daugherty, Betsy  
; APPLICANT: Rogers, No. US20020123618A1ma  
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and  
; FILE REFERENCE: 00-1229  
; CURRENT APPLICATION NUMBER: US/09/928,175  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/224,455  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 21  
LENGTH: 718  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-928-175-21

Query Match 49.5%; Score 409; DB 10; Length 718;

Best Local Similarity 58.9%; Pred. No. 1e-36;  
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;

QY 4 YSAVAFGLGINLAFAITIVFSYSGMFYSVHOSAITATIRNOVKKEMILAKRFFPIVPTDA 63  
Db 553 YSLGIFLGIVNLAFIVIVISYVTWFCSHKXTALQTAVERSHIGEVAVANRFFPIVPSDA 612  
QY 64 LCMPIFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRFMHN 123  
Db 613 ICMPIFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRF 672  
QY 124 YROR 127  
Db 673 HRRK 676

RESULT 7  
US-09-928-175-20  
; Sequence 20, Application US/09928175  
; Patent No. US20020123618A1  
; GENERAL INFORMATION:  
; APPLICANT: Paszty, Christopher J.  
; APPLICANT: Gong, Jianhua  
; APPLICANT: Daugherty, Betsy  
; APPLICANT: Rogers, No. US20020123618A1ma  
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and  
; FILE REFERENCE: 00-1229  
; CURRENT APPLICATION NUMBER: US/09/928,175  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/224,455  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 737  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-928-175-20

Query Match 49.5%; Score 409; DB 10; Length 737;  
Best Local Similarity 58.9%; Pred. No. 1.1e-36;  
Matches 73; Conservative 27; Mismatches 24; Indels 0; Gaps 0;  
QY 4 YSAVAFGLGINLAFAITIVFSYSGMFYSVHOSAITATIRNOVKKEMILAKRFFPIVPTDA 63  
Db 572 YSLGIFLGIVNLAFIVIVISYVTWFCSHKXTALQTAVERSHIGEVAVANRFFPIVPSDA 631  
QY 64 LCMPIFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRFMHN 123  
Db 632 ICMPIFVAKPLSLQVEIPGTTISWVVGYSAINSAIPLIYLTTRPFKEMTHRF 691  
QY 124 YROR 127  
Db 692 HRRK 695

RESULT 8  
US-09-928-175-13  
; Sequence 13, Application US/09928175  
; Patent No. US20020123618A1  
; GENERAL INFORMATION:  
; APPLICANT: Paszty, Christopher J.  
; APPLICANT: Gong, Jianhua  
; APPLICANT: Daugherty, Betsy  
; APPLICANT: Rogers, No. US20020123618A1ma  
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and  
; FILE REFERENCE: 00-1229  
; CURRENT APPLICATION NUMBER: US/09/928,175  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/224,455  
; PRIOR FILING DATE: 2000-08-10



```
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-13
```

```
Query Match
Best Local Similarity 47.5%; Score 392.5; DB 10; Length 646;
Best Local Similarity 59.1%; Pred. No. 5.7e-35;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 4 YSVAIFGIMLAFFIIVSGSMFVSHOSAITATEIRNOVKKEMILAKRFFFIPTDA 63
DB 481 YSLGIFLGVNLAFILIVFSYITMFCSTOKTALQTEVRNCGREVAVANRFFFIPTSDA 540
QY 64 LCMWIPFVAKPLSLQVEIPGTTTSWVVIIGYSAINSLNPILYTLTTRPFKEMIHREFWHN 123
DB 541 ICWIPFVVKIISLFRVEIPDTMTSMIVIFFLPVNSALNPILYTLTTRPFKDKLQKLLHK 600
QY 124 YRORKSM 130
DB 601 H-QRKSI 606
```

## RESULT 9

```
US-09-928-175-12
; Sequence 12, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 682
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-12
```

```
Query Match
Best Local Similarity 47.5%; Score 392.5; DB 10; Length 682;
Best Local Similarity 59.1%; Pred. No. 6.1e-35;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 4 YSVAIFGIMLAFFIIVSGSMFVSHOSAITATEIRNOVKKEMILAKRFFFIPTDA 63
DB 517 YSLGIFLGVNLAFILIVFSYITMFCSTOKTALQTEVRNCGREVAVANRFFFIPTSDA 576
QY 64 LCMWIPFVAKPLSLQVEIPGTTTSWVVIIGYSAINSLNPILYTLTTRPFKEMIHREFWHN 123
DB 577 ICWIPFVVKIISLFRVEIPDTMTSMIVIFFLPVNSALNPILYTLTTRPFKDKLQKLLHK 636
QY 124 YRORKSM 130
DB 637 H-QRKSI 642
```

```
RESULT 10
US-09-928-175-8
; Sequence 8, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
```

```
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-8
```

```
Query Match
Best Local Similarity 47.5%; Score 392.5; DB 10; Length 694;
Best Local Similarity 59.1%; Pred. No. 6.2e-35;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 4 YSVAIFGIMLAFFIIVSGSMFVSHOSAITATEIRNOVKKEMILAKRFFFIPTDA 63
DB 529 YSLGIFLGVNLAFILIVFSYITMFCSTOKTALQTEVRNCGREVAVANRFFFIPTSDA 588
QY 64 LCMWIPFVAKPLSLQVEIPGTTTSWVVIIGYSAINSLNPILYTLTTRPFKEMIHREFWHN 123
DB 589 ICWIPFVVKIISLFRVEIPDTMTSMIVIFFLPVNSALNPILYTLTTRPFKDKLQKLLHK 648
QY 124 YRORKSM 130
DB 649 H-QRKSI 654
```

## RESULT 11

```
US-09-928-175-3
; Sequence 3, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-928-175-3
```

```
Query Match
Best Local Similarity 47.5%; Score 392.5; DB 10; Length 718;
Best Local Similarity 59.1%; Pred. No. 6.5e-35;
Matches 75; Conservative 25; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 4 YSVAIFGIMLAFFIIVSGSMFVSHOSAITATEIRNOVKKEMILAKRFFFIPTDA 63
DB 553 YSLGIFLGVNLAFILIVFSYITMFCSTOKTALQTEVRNCGREVAVANRFFFIPTSDA 612
QY 64 LCMWIPFVAKPLSLQVEIPGTTTSWVVIIGYSAINSLNPILYTLTTRPFKEMIHREFWHN 123
DB 613 ICWIPFVVKIISLFRVEIPDTMTSMIVIFFLPVNSALNPILYTLTTRPFKDKLQKLLHK 672
QY 124 YRORKSM 130
DB 673 H-QRKSI 678
```





